

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2004, 08:40:22 ; Search time 34.3333 Seconds  
(without alignments)  
82.295 Million cell updates/sec

Title: US-09-529-206d-4\_COPY\_127\_136

Perfect score: 47

Sequence: 1 TVSGNLTIR 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	47	100.0	10	2	AAY05980	Human can	
2	47	100.0	10	2	AAY05988	Human can	
3	47	100.0	14	4	AAE07788	Human NY	
4	47	100.0	14	4	AAE07728	Human NY	
5	47	100.0	15	4	AAE07726	Human NY	
6	47	100.0	15	4	AAE07727	Human NY	
7	47	100.0	15	4	AAE07786	Human NY	
8	47	100.0	15	4	AAE07787	Human NY	
9	47	100.0	18	3	AAI52440	Human tum	
10	47	100.0	18	4	AAB69940	Human NY	
11	47	100.0	18	4	AAB69944	Human NY	
12	47	100.0	18	4	AAU01544	Human NY	
13	47	100.0	18	4	AAU01540	HLA-DR53	
14	47	100.0	18	4	AAE07769	Human NY	
15	47	100.0	20	4	AAE07742	Human ESO	
16	47	100.0	25	4	AAE07718	Human NY	
17	47	100.0	25	7	ADD71521	HLA-DP4 b	
18	47	100.0	27	4	AAE07717	Human NY	
19	47	100.0	30	5	AAU85110	Human NYN	
20	47	100.0	36	5	ABG79131	Human NYN	
21	47	100.0	180	2	AAW62584	Cancer as	
22	47	100.0	180	2	AAW69665	Human NY	
23	47	100.0	180	2	AAU05965	Human can	
24	47	100.0	180	3	AAY52430	Human tum	
25	47	100.0	180	3	AAY70862	Human tum	

26	47	100.0	180	3	AAB03154	Human oes
27	47	100.0	180	4	AAB69946	Human NY
28	47	100.0	180	4	AAG67164	Amino aci
29	47	100.0	180	4	AAU01535	Human NY
30	47	100.0	180	4	AAE07714	Human NY
31	47	100.0	180	5	AAU84818	Human NYN
32	47	100.0	180	5	AAU11543	Human tum
33	47	100.0	180	6	ABR58672	Human can
34	47	100.0	180	6	ABR48210	Human bla
35	47	100.0	180	6	ABU56508	Lung canc
36	47	100.0	180	6	ABU56694	Lung canc
37	47	100.0	180	6	ABP74198	Human NY
38	47	100.0	180	6	ABU64816	Human NY
39	47	100.0	180	6	ABR83438	Human NY
40	47	100.0	180	7	ADC09576	NY-ESO-1
41	47	100.0	180	7	ADD35568	Human NY
42	47	100.0	180	7	ADD25510	Binding d
43	47	100.0	397	4	AAE13122	NY-ESO-1C
44	47	100.0	3541	5	AAU85130	Human mel
45	42	89.4	9	2	AAY06030	Human can

#### ALIGNMENTS

RESULT 1  
AAY05980  
ID AAY05980 standard; peptide; 10 AA.  
XX  
AC AAY05980;  
XX  
DT 16-AUG-1999 (first entry)  
XX  
DE Human cancer antigen NY ESO-1/CAG-3 ORF1 cancer peptide ESO10-127.  
XX  
KW NY ESO-1/CAG-3 gene; cancer peptide; antigen; human;  
KW leukemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
KW vaccine.  
XX  
OS Homo sapiens.  
XX  
PN WO9918206-A2.  
XX  
PD 15-APR-1999.  
XX  
PF 21-SEP-1998; 98WO-US019609.  
XX  
PR 08-OCT-1997; 97US-0061428P.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Wang RF, Rosenberg SA;  
XX  
DR WPI; 1999-277270/23.  
XX  
PT Cancer antigen NY ESO1/CAG-3.  
XX  
PS Claim 17; Page 64; 88pp; English.  
XX  
CC This sequence represents cancer peptide ESO10-127 that corresponds to  
CC amino acid residues 127-136 of human ESO-1/CAG-3 (or CAG-3) ORF1 (see  
CC AAY05965), a new and potent tumour antigen capable of eliciting an  
CC antigen specific immune response by T cells. Cancer peptides derived from  
CC CAG-3 ORF1, CAG-3 ORF2 (see AAY05966), portions of them and their  
CC variants (see AAY05967-87), are useful as cancer vaccines that protect  
CC against cancer. The invention provides: vectors and host cells (also  
CC useful as vaccines); a method of diagnosis of cancer or precancer; a  
CC transgenic animal; antisense oligonucleotides that inhibit expression of  
CC the cancer peptide or tumour antigen; antibodies reacting with a CAG-3

CC cancer peptide, useful in diagnostic and detection assays; and methods  
 CC for preventing or inhibiting cancer by administering a cancer peptide,  
 CC with or without an HLA molecule. The cancer peptides form part of, or are  
 CC derived from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine cancer,  
 CC cervical cancer, bladder cancer, kidney cancer and adenocarcinomas such  
 CC as breast, prostate, ovarian, pancreatic and thyroid cancers. Melanoma is  
 CC treated by inducing cancer-specific T cells in vitro for subsequent  
 CC return to a patient  
 XX  
 XX Sequence 10 AA;

Query Match 100.0%; Score 47; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0029;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10  
 |||||  
 Db 1 TVSGNLTIR 10

RESULT 2  
 AAY05988  
 ID AAY05988 standard; peptide; 10 AA.  
 XX  
 AC AAY05988;  
 XX  
 DT 16-AUG-1999 (first entry)  
 XX  
 DE Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 XX vaccine; human leukocyte antigen; HLA.

OS Homo sapiens.

XX

XX WO99:18206-A2.

XX 15-APR-1999.

XX 21-SEP-1998; 98WO-US019609.

XX 08-OCT-1997; 97US-0061428P.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Wang RP, Rosenberg SA;

XX WPI; 1999-27270/23.

XX Cancer antigen NY ESO1/CAG-3.

XX Example 10; Page 42; 88pp; English.

XX This peptide was identified as an HLA peptide motif following a screen  
 CC for epitopes from the coding region of human NY ESO-1/CAG-3 ORF1 (see  
 CC AAX9599). 30 Epitopes (see AAY05988-Y06017) were identified. The present  
 CC peptide (ranked 1) corresponds to amino acid residues 127-136 of CAG-1  
 CC ORF1 (see AAY05963). CAG-1 is a new and potent tumour antigen capable of  
 CC eliciting an antigen specific immune response by T cells. Cancer peptides  
 CC (see AAY05967-87) derived from CAG-3, portions of CAG-3 and their  
 CC variants, are useful as cancer vaccines. A claimed method of preventing  
 CC or inhibiting cancer involves administering a cancer peptide, with or  
 CC without an HLA molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma, lymphoma,  
 CC sarcoma, lung cancer, liver cancer, leukaemia, uterine cancer, cervical  
 CC cancer, bladder cancer, kidney cancer and adenocarcinomas such as breast,  
 CC prostate, ovarian, pancreatic and thyroid cancers

XX Sequence 10 AA;

Query Match 100.0%; Score 47; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0029;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10  
 |||||  
 Db 1 TVSGNLTIR 10

RESULT 3  
 AAE07788  
 ID AAE07788 standard; peptide; 14 AA.

XX AAE07788;

XX 06-NOV-2001 (first entry)

XX Human NY ESO-1 peptide #22 to characterise epitope recognised by TE4-1.

XX Human; major histocompatibility complex; MHC; vaccine; metastasis;  
 KW class II restricted T cell epitope; MHC-II epitope; cancer antigen;  
 KW NY ESO-1 protein; CD4+ T lymphocyte; human leukocyte antigen; HLA;  
 KW tumour-specific humoral-mediated immunity; cancer; cytostatic;  
 KW immunotherapy.

XX Homo sapiens.

XX WO200155393-A2.

XX 02-AUG-2001.

XX 26-JAN-2001; 2001WO-US002765.

XX 28-JAN-2000; 2000US-0179004P.

XX 29-SEP-2000; 2000US-0237107P.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Wang R, Rosenberg SA, Zeng G;

XX WPI; 2001-496851/54.

XX New NY-ESO cancer peptide or MHC class II restricted T cell epitopes,  
 PT useful as immunogen and vaccine for inhibiting cancer in a mammal or as  
 PT protection from metastasis.

XX Example 6; Fig 6A; 134pp; English.

XX The invention relates to the identification and isolation of major  
 CC histocompatibility (MHC) class II restricted T cell epitope (MHC-II  
 CC epitope) derived from the cancer antigen, NY ESO-1. The MHC-II epitopes  
 CC from NY ESO-1 are recognised by CD4+ T lymphocytes in an human leucocyte  
 CC antigen (HLA) class II restricted manner, in particular HLA-DR or HLA-DP  
 CC restricted. The products of the gene are promising candidates for  
 CC immunotherapeutic strategies for the prevention, treatment and diagnosis  
 CC of patients with cancer. The cancer epitopes are useful as immunogen and  
 CC vaccine to inhibit or to prevent cancer in a mammal by eliciting CD4+ T  
 CC lymphocytes resulting in protection of the recipient from development of  
 CC cancer and protection from metastasis, or by inhibiting the growth of  
 CC cells expressing the NY-ESO-1 gene product. The cancer peptides are also  
 CC useful as diagnostic agent to detect the presence of cancer, to enhance  
 CC the generation of antibody and/or CD8+ T cell responses against any given  
 CC target antigen and/or hapten and to induce tumour-specific humoral-  
 CC mediated immunity against cancer. The present sequence is human NY ESO-1  
 CC peptide used in the characterisation of the NY ESO-1 epitope recognised  
 CC by TE4-1

XX Sequence 14 AA;

Query Match 100.0%; Score 47; DB 4; Length 14;

Best Local Similarity 100.0%; Pred. No. 0.0045; Mismatches 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0

QY 1 TVSGNLTIR 10  
|||||||  
Db 4 TVSGNLTIR 13

RESULT 4  
AAE07728  
ID AAE07728 standard; peptide; 14 AA.

AC AAE07728;  
XX  
XX 06-NOV-2001 (first entry)  
XX  
XX Human NY ESO-1 MHC class II restricted T cell epitope #14.  
XX  
XX Human; major histocompatibility complex; MHC; vaccine; metastasis;  
KW class II restricted T cell epitope; MHC-II epitope; cancer antigen;  
KW NY ESO-1 protein; CD4+ T lymphocyte; human leucocyte antigen; HLA;  
KW tumour-specific humoral-mediated immunity; cancer; cytostatic;  
KW immunotherapy.

XX Homo sapiens.

OS WO200155393-A2.

EN 02-AUG-2001.

XX 26-JAN-2001; 2001WO-US002765.

XX 28-JAN-2000; 2000US-0179004P.

PR 29-SEP-2000; 2000US-0237107P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA Wang R, Rosenberg SA, Zeng G;

PI WPI; 2001-496851/54.

XX New NY-ESO cancer peptide or MHC class II restricted T cell epitopes,  
PT useful as immunogen and vaccine for inhibiting cancer in a mammal or as  
PT protection from metastasis.

XX Claim 4; Page 16; 134pp; English.

XX The invention relates to the identification and isolation of major  
CC histocompatibility (MHC) class II restricted T cell epitope (MHC-II  
CC epitope) derived from the cancer antigen, NY ESO-1. The MHC-II epitopes  
CC from NY ESO-1 are recognised by CD4+ T lymphocytes in an human leucocyte  
CC antigen (HLA) class II restricted manner, in particular HLA-DR or HLA-DP  
CC restricted. The products of the gene are promising candidates for  
CC immunotherapeutic strategies for the prevention, treatment and diagnosis  
CC of patients with cancer. The cancer epitopes are useful as immunogen and  
CC vaccine to inhibit or to prevent cancer in a mammal by eliciting CD4+ T  
CC lymphocytes resulting in protection of the recipient from development of  
CC cancer and protection from metastasis, or by inhibiting the growth of  
CC cells expressing the NY-ESO-1 gene product. The cancer peptides are also  
CC useful as diagnostic agent to detect the presence of cancer, to enhance  
CC the generation of antibody and/or CD8+ T cell responses against any given  
CC target antigen and/or hapten and to induce tumour-specific humoral-  
CC mediated immunity against cancer. The present sequence is MHC class II  
CC restricted T cell epitope of human NY ESO-1 protein

XX Sequence 14 AA;

Query Match 100.0%; Score 47; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.0045;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10  
|||||||

Db 4 TVSGNLTIR 13

RESULT 5  
AAE07726

ID AAE07726 standard; peptide; 15 AA.

XX AAE07726;

XX 06-NOV-2001 (first entry)

XX Human NY ESO-1 MHC class II restricted T cell epitope #12.

XX Human; major histocompatibility complex; MHC; vaccine; metastasis;  
KW class II restricted T cell epitope; MHC-II epitope; cancer antigen;  
KW NY ESO-1 protein; CD4+ T lymphocyte; human leucocyte antigen; HLA;  
KW tumour-specific humoral-mediated immunity; cancer; cytostatic;  
KW immunotherapy.

XX Homo sapiens.

XX WO200155393-A2.

XX 02-AUG-2001.

XX 26-JAN-2001; 2001WO-US002765.

XX 28-JAN-2000; 2000US-0179004P.

PR 29-SEP-2000; 2000US-0237107P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA Wang R, Rosenberg SA, Zeng G;

PI WPI; 2001-496851/54.

XX New NY-ESO cancer peptide or MHC class II restricted T cell epitopes,  
PT useful as immunogen and vaccine for inhibiting cancer in a mammal or as  
PT protection from metastasis.

XX Claim 4; Page 16; 134pp; English.

XX The invention relates to the identification and isolation of major  
CC histocompatibility (MHC) class II restricted T cell epitope (MHC-II  
CC epitope) derived from the cancer antigen, NY ESO-1. The MHC-II epitopes  
CC from NY ESO-1 are recognised by CD4+ T lymphocytes in an human leucocyte  
CC antigen (HLA) class II restricted manner, in particular HLA-DR or HLA-DP  
CC restricted. The products of the gene are promising candidates for  
CC immunotherapeutic strategies for the prevention, treatment and diagnosis  
CC of patients with cancer. The cancer epitopes are useful as immunogen and  
CC vaccine to inhibit or to prevent cancer in a mammal by eliciting CD4+ T  
CC lymphocytes resulting in protection of the recipient from development of  
CC cancer and protection from metastasis, or by inhibiting the growth of  
CC cells expressing the NY-ESO-1 gene product. The cancer peptides are also  
CC useful as diagnostic agent to detect the presence of cancer, to enhance  
CC the generation of antibody and/or CD8+ T cell responses against any given  
CC target antigen and/or hapten and to induce tumour-specific humoral-  
CC mediated immunity against cancer. The present sequence is MHC class II  
CC restricted T cell epitope of human NY ESO-1 protein

XX Sequence 15 AA;

Query Match 100.0%; Score 47; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0049;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10  
|||||||  
Db 6 TVSGNLTIR 15

RESULT 6  
AAE07727

ID AAE07727 standard; peptide; 15 AA.  
 AC AAE07727;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human NY ESO-1 MHC class II restricted T cell epitope #13.  
 XX  
 KW Human; major histocompatibility complex; MHC; vaccine; metastasis;  
 KW class II restricted T cell epitope; MHC-II epitope; cancer antigen;  
 KW NY ESO-1 protein; CD4+ T lymphocyte; human leucocyte antigen; HLA;  
 KW tumour-specific humoral-mediated immunity; cancer; cytostatic;  
 KW immunotherapy.  
 XX  
 OS Homo sapiens.  
 OS  
 XX WO200155393-A2.  
 PN  
 XX 02-AUG-2001.  
 PD  
 XX 26-JAN-2001; 2001WO-US002765.  
 PF  
 XX 28-JAN-2000; 2000US-0179004P.  
 PR  
 XX 29-SEP-2000; 2000US-0237107P.  
 PR  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 XX Wang R, Rosenberg SA, Zeng G;  
 PI  
 XX WPI; 2001-496851/54.  
 DR  
 XX  
 XX New NY-ESO cancer peptide or MHC class II restricted T cell epitopes,  
 XX useful as immunogen and vaccine for inhibiting cancer in a mammal or as  
 XX protection from metastasis.  
 XX  
 PS Claim 4; Page 16; 134pp; English.  
 XX  
 XX The invention relates to the identification and isolation of major  
 XX histocompatibility (MHC) class II restricted T cell epitope (MHC-II  
 XX epitope) derived from the cancer antigen, NY ESO-1. The MHC-II epitopes  
 XX from NY ESO-1 are recognised by CD4+ T lymphocytes in an human leucocyte  
 XX antigen (HLA) class II restricted manner, in particular HLA-DR or HLA-DP  
 XX restricted. The products of the gene are promising candidates for  
 XX immunotherapeutic strategies for the prevention, treatment and diagnosis  
 XX of patients with cancer. The cancer epitopes are useful as immunogen and  
 XX vaccine to inhibit or to prevent cancer in a mammal by eliciting CD4+ T  
 XX lymphocytes resulting in protection of the recipient from development of  
 XX cancer and protection from metastasis, or by inhibiting the growth of  
 XX cells expressing the NY-ESO-1 gene product. The cancer peptides are also  
 XX useful as diagnostic agent to detect the presence of cancer, to enhance  
 XX the generation of antibody and/or CD8+ T cell responses against any given  
 XX target antigen and/or hapten and to induce tumour-specific humoral-  
 XX mediated immunity against cancer. The present sequence is MHC class II  
 XX restricted T cell epitope of human NY ESO-1 protein  
 XX  
 SQ Sequence 15 AA;  
 Query Match 100.0%; Score 47; DB 4; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.0049;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TVSGNLTIR 10  
 DB 5 TVSGNLTIR 14  
 RESULT 7  
 AAE07786  
 ID AAE07786 standard; peptide; 15 AA.  
 XX  
 AC AAE07786;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX

XX Human NY ESO-1 peptide #20 to characterise epitope recognised by TE4-1.  
 DE  
 XX  
 KW Human; major histocompatibility complex; MHC; vaccine; metastasis;  
 KW class II restricted T cell epitope; MHC-II epitope; cancer antigen;  
 KW NY ESO-1 protein; CD4+ T lymphocyte; human leucocyte antigen; HLA;  
 KW tumour-specific humoral-mediated immunity; cancer; cytostatic;  
 KW immunotherapy.  
 XX  
 OS Homo sapiens.  
 OS  
 XX WO200155393-A2.  
 PN  
 XX 02-AUG-2001.  
 PD  
 XX 26-JAN-2001; 2001WO-US002765.  
 PF  
 XX 28-JAN-2000; 2000US-0179004P.  
 PR  
 XX 29-SEP-2000; 2000US-0237107P.  
 PR  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 XX Wang R, Rosenberg SA, Zeng G;  
 PI  
 XX WPI; 2001-496851/54.  
 DR  
 XX  
 XX New NY-ESO cancer peptide or MHC class II restricted T cell epitopes,  
 XX useful as immunogen and vaccine for inhibiting cancer in a mammal or as  
 XX protection from metastasis.  
 XX  
 PS Example 6; Fig 6A; 134pp; English.  
 XX  
 XX The invention relates to the identification and isolation of major  
 XX histocompatibility (MHC) class II restricted T cell epitope (MHC-II  
 XX epitope) derived from the cancer antigen, NY ESO-1. The MHC-II epitopes  
 XX from NY ESO-1 are recognised by CD4+ T lymphocytes in an human leucocyte  
 XX antigen (HLA) class II restricted manner, in particular HLA-DR or HLA-DP  
 XX restricted. The products of the gene are promising candidates for  
 XX immunotherapeutic strategies for the prevention, treatment and diagnosis  
 XX of patients with cancer. The cancer epitopes are useful as immunogen and  
 XX vaccine to inhibit or to prevent cancer in a mammal by eliciting CD4+ T  
 XX lymphocytes resulting in protection of the recipient from development of  
 XX cancer and protection from metastasis, or by inhibiting the growth of  
 XX cells expressing the NY-ESO-1 gene product. The cancer peptides are also  
 XX useful as diagnostic agent to detect the presence of cancer, to enhance  
 XX the generation of antibody and/or CD8+ T cell responses against any given  
 XX target antigen and/or hapten and to induce tumour-specific humoral-  
 XX mediated immunity against cancer. The present sequence is human NY ESO-1  
 XX peptide used in the characterisation of the NY ESO-1 epitope recognised  
 XX by TE4-1  
 XX  
 SQ Sequence 15 AA;  
 Query Match 100.0%; Score 47; DB 4; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.0049;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TVSGNLTIR 10  
 DB 6 TVSGNLTIR 15  
 RESULT 8  
 AAE07787  
 ID AAE07787 standard; peptide; 15 AA.  
 XX  
 AC AAE07787;  
 XX  
 DT 06-NOV-2001 (first entry)  
 DT  
 XX Human NY ESO-1 peptide #21 to characterise epitope recognised by TE4-1.  
 DE  
 XX Human; major histocompatibility complex; MHC; vaccine; metastasis;  
 KW

KW class II restricted T cell epitope; MHC-II epitope; cancer antigen;  
 KW NY ESO-1 protein; CD4+ T lymphocyte; human leukocyte antigen; HLA;  
 KW tumour-specific humoral-mediated immunity; cancer; cytostatic;  
 XX immunotherapy.  
 XX  
 XX Homo sapiens.  
 XX  
 XX WO200155393-A2.  
 XX  
 XX PD 02-AUG-2001.  
 XX  
 XX PF 26-JAN-2001; 2001WO-US002765.  
 XX  
 XX XX 28-JAN-2000; 2000US-0179004P.  
 XX PR 29-SEP-2000; 2000US-0237107P.  
 XX  
 XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 XX XX Wang R, Rosenberg SA, Zeng G;  
 FI WPI; 2001-496851/54.  
 XX  
 XX DR New NY-ESO cancer peptide or MHC class II restricted T cell epitopes,  
 PT useful as immunogen and vaccine for inhibiting cancer in a mammal or as  
 PT protection from metastasis.  
 XX  
 XX PS Example 6; Fig 6A; 134pp; English.  
 XX  
 CC The invention relates to the identification and isolation of major  
 CC histocompatibility (MHC) class II restricted T cell epitope (MHC-II  
 CC epitope) derived from the cancer antigen, NY ESO-1. The MHC-II epitopes  
 CC from NY ESO-1 are recognised by CD4+ T lymphocytes in an human leukocyte  
 CC antigen (HLA) class II restricted manner, in particular HLA-DR or HLA-DP  
 CC restricted. The products of the gene are promising candidates for  
 CC immunotherapeutic strategies for the prevention, treatment and diagnosis  
 CC of patients with cancer. The cancer epitopes are useful as immunogen and  
 CC vaccine to inhibit or to prevent cancer in a mammal by eliciting CD4+ T  
 CC lymphocytes resulting in protection of the recipient from development of  
 CC cancer and protection from metastasis, or by inhibiting the growth of  
 CC cells expressing the NY-ESO-1 gene product. The cancer peptides are also  
 CC useful as diagnostic agent to detect the presence of cancer, to enhance  
 CC the generation of antibody and/or CD8+ T cell responses against any given  
 CC target antigen and/or to haptent and to induce tumour-specific humoral-  
 CC mediated immunity against cancer. The present sequence is human NY ESO-1  
 CC peptide used in the characterisation of the NY ESO-1 epitope recognised  
 CC by TE4-1  
 XX  
 XX SQ Sequence 15 AA;

Query Match 100.0%; Score 47; DB 4; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.0049;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10  
 |||||  
 Db 5 TVSGNLTIR 14

RESULT 9  
 AAY52440  
 ID AAY52440 standard; protein; 18 AA.  
 XX  
 XX AC AAY52440;  
 XX  
 XX DT 15-FEB-2000 (first entry)  
 XX  
 XX DE Human tumour antigen NY-ESO-1 peptide #13.  
 XX  
 KW Cancer; tumour; antigen; MHC; major histocompatibility complex; Class II;  
 KW T-cell; helper; stimulation; proliferation; treatment; diagnosis;  
 KW prevention; melanoma; breast cancer; ovarian cancer; prostate cancer;  
 KW hepatoma; thyroid cancer; bladder cancer; lung cancer; lymphoma.  
 XX

OS Synthetic.  
 OS Homo sapiens.  
 XX  
 XX WO9953938-A1.  
 XX  
 XX PD 28-OCT-1999.  
 XX  
 XX PF 24-MAR-1999; 99WO-US006875.  
 XX  
 XX PR 17-APR-1998; 98US-00062422.  
 XX PR 02-OCT-1998; 98US-00165546.  
 XX  
 XX PA (LUDW-) LUDWIG INST CANCER RES.  
 XX  
 XX XX Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;  
 FI Gure A, Rittner G;  
 XX  
 XX DR WPI; 2000-038483/03.  
 XX  
 XX PT Novel peptides which bind to MHC class I and MHC class II molecules,  
 PT useful for therapeutic and diagnostic purposes.  
 XX  
 XX PS Claim 4; Page 22; 49pp; English.  
 XX  
 CC Peptides #8-#13 (AAY52435-Y52440) are peptides derived from the human  
 CC tumour antigen, NY-ESO-1 (AAY52430) which can bind to MHC(major  
 CC histocompatibility Class II HLA-DR53 molecules, thereby stimulating  
 CC proliferation of helper T-cells. cDNA encoding NY-ESO-1 was initially  
 CC isolated from an oesophagus squamous cell cancer cDNA library. Tissue  
 CC localisation studies revealed it to be expressed at high levels in normal  
 CC ovary and testis but not in normal colon, kidney, liver, brain,  
 CC oesophagus and skin. It was expressed in certain tumours and tumour cell  
 CC lines with some degree of frequency - these included melanoma specimens  
 CC and cell lines, and breast and bladder cancer specimens, with expression  
 CC in other tumour types being sporadic. These NY-ESO-1-derived peptides may  
 CC be used in methods and compositions used for the treatment, diagnosis and  
 CC prevention of cancers (such as melanoma, breast cancer, prostate cancer,  
 CC lung cancer, hepatoma, ovarian cancer, thyroid cancer, bladder cancer, or  
 CC lymphoma) and to stimulate the proliferation of T cells  
 XX  
 XX SQ Sequence 18 AA;

Query Match 100.0%; Score 47; DB 3; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 0.0062;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10  
 |||||  
 Db 1 TVSGNLTIR 10

RESULT 10  
 AAB69940  
 ID AAB69940 standard; peptide; 18 AA.  
 XX  
 XX AC AAB69940;  
 XX  
 XX DT 27-APR-2001 (first entry)  
 XX  
 XX DE Human NY-ESO-1 HLA-DR53 binding motif #2.  
 XX  
 KW Human; NY-ESO-1; HLA; human leukocyte antigen; CTL; cytotoxic T cell;  
 KW HLA-A2; HLA-DR53; melanoma; adenocarcinoma; bladder carcinoma;  
 KW non-small cell lung carcinoma; tumour status determination.  
 XX  
 XX OS Homo sapiens.  
 XX  
 XX PN WO200107917-A1.  
 XX  
 XX PD 01-FEB-2001.  
 XX  
 XX PF 14-JUL-2000; 2000WO-US019220.  
 XX

PR 23-JUL-1999; 99US-00359503.  
XX (LUDW-) LUDWIG INST CANCER RES.  
PA (SLOK ) SLOAN KETTERING INST CANCER RES.  
PA (CORR ) CORNELL RES FOUND INC.  
XX  
XX Jager E, Stockert E, Old LJ, Knuth A, Chen Y, Scanlan M;  
XX WPI; 2001-182822/18.  
XX  
XX Method useful for determining the status (e.g. progression, regression or  
PT stability of the disease) of a cancerous condition, involves determining  
PT the levels of NY-ESO-1 specific antibodies in a sample taken from a  
PT patient.  
XX  
XX Example 16; Page 27; 50pp; English.  
PS  
CC The present sequence is given in a specification relating to a method for  
CC determining the status of a cancerous condition in a patient with a  
CC tumour that expresses NY-ESO-1. The method comprises assaying a sample  
CC taken from the patient for antibodies that specifically bind to the NY-  
CC ESO-1 and comparing the value obtained to a prior value obtained from  
CC assay of a prior sample taken from the patient. Any difference between  
CC the values is indicative of a change in status of the cancerous  
CC condition. The method is useful for determining whether a cancerous  
CC condition is progressing, regressing or remaining stable, in particular  
CC in patients receiving treatment for a melanoma, adenocarcinoma, non-small  
CC cell lung carcinoma or bladder carcinoma  
XX  
XX Sequence 18 AA;  
SQ  
Query Match 100.0%; Score 47; DB 4; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.0062;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TVSGNLTIR 10  
DB |||||  
1 TVSGNLTIR 10  
RESULT 12  
AAU01544  
ID AAU01544 standard; peptide; 18 AA.  
XX  
AC AAU01544;  
XX  
DT 18-JUL-2001 (first entry)  
XX  
DE HLA-DR53 recognising NY-ESO-1 peptide #6.  
XX  
KW NY-ESO-1; human; tumour rejection antigen precursor; SSX-2; MHC Class II;  
KW major histocompatibility complex; helper T cell; HLA-DR; cancer;  
KW human leukocyte antigen-determining region; disease progression;  
KW disease regression; disease onset; body tissue; body fluid; enzyme label;  
KW radioactive label; monoclonal antibody.  
XX  
OS Homo sapiens.  
XX  
FN WO200123560-A2.  
XX  
PD 05-APR-2001.  
XX  
PF 26-SEP-2000; 2000WO-US026411.  
XX  
PR 29-SEP-1999; 99US-00408036.  
XX  
PA (LUDW-) LUDWIG INST CANCER RES.  
XX  
PI Tureci O, Sahin U, Pfreundschuh M;  
XX WPI; 2001-266156/27.  
XX  
DR Polypeptides binding to major histocompatibility complex class II human  
PT leukocyte antigen-determining region molecule having amino acid sequence  
PT found in tumor rejection antigen precursor used for stimulating  
PT proliferation of helper T cells.  
XX  
XX Example 13; Page 19; 62pp; English.  
PS  
XX The sequence represents a human NY-ESO-1 tumour rejection antigen  
CC precursor fragment which recognises and binds to HLA-DR53. NY-ESO-1 and  
CC SSX-2 polypeptides, or fragments of, bind to major histocompatibility  
CC complex (MHC) Class II molecules such as human leukocyte antigen-  
CC determining region (HLA-DR) molecules and stimulate proliferation of

PR 23-JUL-1999; 99US-00359503.  
XX (LUDW-) LUDWIG INST CANCER RES.  
PA (SLOK ) SLOAN KETTERING INST CANCER RES.  
PA (CORR ) CORNELL RES FOUND INC.  
XX  
XX Jager E, Stockert E, Old LJ, Knuth A, Chen Y, Scanlan M;  
XX WPI; 2001-182822/18.  
XX  
XX Method useful for determining the status (e.g. progression, regression or  
PT stability of the disease) of a cancerous condition, involves determining  
PT the levels of NY-ESO-1 specific antibodies in a sample taken from a  
PT patient.  
XX  
XX Example 16; Page 27; 50pp; English.  
PS  
CC The present sequence is given in a specification relating to a method for  
CC determining the status of a cancerous condition in a patient with a  
CC tumour that expresses NY-ESO-1. The method comprises assaying a sample  
CC taken from the patient for antibodies that specifically bind to the NY-  
CC ESO-1 and comparing the value obtained to a prior value obtained from  
CC assay of a prior sample taken from the patient. Any difference between  
CC the values is indicative of a change in status of the cancerous  
CC condition. The method is useful for determining whether a cancerous  
CC condition is progressing, regressing or remaining stable, in particular  
CC in patients receiving treatment for a melanoma, adenocarcinoma, non-small  
CC cell lung carcinoma or bladder carcinoma  
XX  
XX Sequence 18 AA;  
SQ  
Query Match 100.0%; Score 47; DB 4; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.0062;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TVSGNLTIR 10  
DB |||||  
7 TVSGNLTIR 16  
RESULT 11  
AAB69944  
ID AAB69944 standard; peptide; 18 AA.  
XX  
AC AAB69944;  
XX  
DT 27-APR-2001 (first entry)  
XX  
DE Human NY-ESO-1 HLA-DR53 binding motif #6.  
XX  
KW Human; NY-ESO-1; HLA; human leukocyte antigen; CTL; cytotoxic T cell;  
KW HLA-A2; HLA-DR53; melanoma; adenocarcinoma; bladder carcinoma;  
KW non-small cell lung carcinoma; tumour status determination.  
XX  
CS Homo sapiens.  
XX  
FN WO200107917-A1.  
XX  
PD 01-FEB-2001.  
XX  
PF 14-JUL-2000; 2000WO-US019220.  
XX  
PR 23-JUL-1999; 99US-00359503.  
XX  
PA (LUDW-) LUDWIG INST CANCER RES.  
PA (SLOK ) SLOAN KETTERING INST CANCER RES.  
PA (CORR ) CORNELL RES FOUND INC.  
XX  
XX Jager E, Stockert E, Old LJ, Knuth A, Chen Y, Scanlan M;  
XX WPI; 2001-182822/18.  
XX  
XX Method useful for determining the status (e.g. progression, regression or



CC the generation of antibody and/or CD8+ T cell responses against any given  
CC target antigen and/or hapten and to induce tumour-specific humoral-  
CC mediated immunity against cancer. The present sequence is human NY ESO-1  
CC HLA DR restricted T cell cancer peptide

XX SQ Sequence 18 AA;  
Query Match 100.0%; Score 47; DB 4; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.0062;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10  
| | | | | | | | | |  
Db 7 TVSGNLTIR 16

RESULT 15  
AAE07742  
ID AAE07742 standard; peptide; 20 AA.

XX AC AAE07742;

XX DT 06-NOV-2001 (first entry)

XX DE Human ESO p126-145 peptide, to identify MHC class II-restricted epitopes.

XX DE Human; major histocompatibility complex; MHC; vaccine; metastasis;  
XX KW class II restricted T cell epitope; MHC-II epitope; cancer antigen;  
XX KW NY ESO-1 protein; CD4+ T lymphocyte; human leucocyte antigen; HLA;  
XX KW tumour-specific humoral-mediated immunity; cancer; cytostatic;  
XX KW immunotherapy.

XX OS Homo sapiens.

XX PN WO200155393-A2.

XX ED 02-AUG-2001.

XX PF 26-JAN-2001; 2001WO-US002765.

XX PR 28-JAN-2000; 2000US-0179004P.

XX PR 29-SEP-2000; 2000US-0237107P.

XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX PI Wang R, Rosenberg SA, Zeng G;

XX DR WPI; 2001-496851/54.

XX PT New NY-ESO cancer peptide or MHC class II restricted T cell epitopes,  
XX PT useful as immunogen and vaccine for inhibiting cancer in a mammal or as  
XX PT protection from metastasis.

XX PS Example 3; Fig 3; 134pp; English.

XX CC The invention relates to the identification and isolation of major  
XX CC histocompatibility (MHC) class II restricted T cell epitope (MHC-II  
XX CC epitope) derived from the cancer antigen, NY ESO-1. The MHC-II epitopes  
XX CC from NY ESO-1 are recognised by CD4+ T lymphocytes in an human leucocyte  
XX CC antigen (HLA) class II restricted manner, in particular HLA-DR or HLA-DP  
XX CC restricted. The products of the gene are promising candidates for  
XX CC immunotherapeutic strategies for the prevention, treatment and diagnosis  
XX CC of patients with cancer. The cancer epitopes are useful as immunogen and  
XX CC vaccine to inhibit or to prevent cancer in a mammal by eliciting CD4+ T  
XX CC lymphocytes resulting in protection of the recipient from development of  
XX CC cancer and protection from metastasis, or by inhibiting the growth of  
XX CC cells expressing the NY-ESO-1 gene product. The cancer peptides are also  
XX CC useful as diagnostic agent to detect the presence of cancer, to enhance  
XX CC the generation of antibody and/or CD8+ T cell responses against any given  
XX CC target antigen and/or hapten and to induce tumour-specific humoral-  
XX CC mediated immunity against cancer. The present sequence is human ESO p126-  
XX CC 145 peptide used in the identification of putative MHC class II -  
XX CC restricted epitopes from HLA-DR4-transgenic mice

XX SQ Sequence 20 AA;

Query Match 100.0%; Score 47; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0071;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10  
| | | | | | | | | |  
Db 2 TVSGNLTIR 11

Search completed: June 22, 2004, 08:51:24  
Job time : 34.3333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2004, 08:49:38 ; Search time 11 Seconds  
(without alignments)

46.933 Million cell updates/sec

Title: US-09-529-206D-4\_COPY\_127\_136

Perfect score: 47

Sequence: 1 TVSGNIIITIR 10

Scoring table: BLOSUM62

Gapop 10.0 ; Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PTCUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	47	100.0	18	3	US-09-359-503-9
2	47	100.0	18	3	US-09-359-503-13
3	47	100.0	180	3	US-08-791-495-9
4	47	100.0	180	3	US-08-937-263B-8
5	47	100.0	180	4	US-09-751-798-8
6	47	100.0	180	4	US-09-392-714-25
7	42	89.4	9	4	US-09-344-040C-125
8	42	89.4	9	4	US-08-833-039A-125
9	38	80.9	9	4	US-09-344-040C-120
10	38	80.9	9	4	US-08-833-039A-120
11	38	80.9	180	2	US-08-791-495-7
12	35	74.5	342	4	US-09-134-001C-5428
13	34	72.3	149	4	US-09-540-236-2614
14	34	72.3	177	4	US-09-732-210-1076
15	33	70.2	97	1	US-08-118-270-329
16	33	70.2	97	5	PCT-US93-08528-329
17	33	70.2	109	4	US-09-341-461-35
18	33	70.2	362	4	US-09-134-001C-5209
19	33	70.2	379	5	PCT-US93-08528-32
20	33	70.2	379	5	PCT-US93-08528-32
21	33	70.2	470	3	US-09-292-071-25
22	33	70.2	470	3	US-09-292-069A-25
23	33	70.2	470	4	US-09-767-013-25
24	33	70.2	470	4	US-09-292-072-25
25	33	70.2	471	1	US-07-817-920-8
26	33	70.2	471	1	US-07-996-772A-11
27	33	70.2	471	1	US-08-370-542-7

Sequence 8, Appli  
Sequence 8, Appli  
Sequence 7, Appli  
Sequence 2, Appli  
Sequence 7, Appli  
Sequence 8, Appli  
Sequence 11, Appli  
Sequence 14, Appli  
Sequence 17, Appli  
Sequence 17, Appli  
Sequence 2, Appli  
Sequence 4, Appli  
Sequence 122, App  
Sequence 228, App  
Sequence 8, Appli  
Sequence 31, Appli  
Sequence 33, Appli

#### ALIGNMENTS

RESULT 1  
US-09-359-503-9  
Sequence 9, Application US/09359503  
Patent No. 6251603  
GENERAL INFORMATION:  
APPLICANT: Jager, Elke; Stockert, Elisabeth; Old, Lloyd J;  
APPLICANT: Knuth, alexander  
TITLE OF INVENTION: Method for Determining Status of A  
TITLE OF INVENTION: Cancerous Condition By Determining Antibodies  
TITLE OF INVENTION: a Patient Sample  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Fulbright & Jaworski, L.L.P.  
STREET: 666 Fifth Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/359,503  
FILING DATE: July 23, 1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/165,546  
FILING DATE: October 2, 1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/062,422  
FILING DATE: April 17, 1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/937,263  
FILING DATE: September 15, 1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/752,182  
FILING DATE: 03-October-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6251603man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5466.1  
TELEPHONE: (212) 318-3000  
TELEFAX: (212) 752-5958  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid

```

; TOPOLOGY: linear
US-09-359-503-9
Query Match      100.0%; Score 47; DB 3; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.008;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10
Db 7 TVSGNLTIR 16

RESULT 2
US-09-359-503-13
; Sequence 13, Application US/09359503
; Patent No. 6251603
; GENERAL INFORMATION:
; APPLICANT: Jager, Elke; Stockert, Elisabeth; Old, Lloyd J;
; APPLICANT: Knuth, Alexander
; TITLE OF INVENTION: Method for Determining Status of A
; TITLE OF INVENTION: Cancerous Condition By Determining Antibodies
; TITLE OF INVENTION: a Patient Sample
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/359,503
; FILING DATE: July 23, 1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/165,546
; FILING DATE: October 2, 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/062,422
; FILING DATE: April 17, 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/937,263
; FILING DATE: September 15, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/752,182
; FILING DATE: 03-October-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6251603man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5466.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-359-503-13
Query Match      100.0%; Score 47; DB 3; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.008;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10
Db 1 TVSGNLTIR 10

us-09-529-206d-4 copy_127_136.ra1 to NY-
; TOPOLOGY: linear
US-08-791-495-9
Query Match      100.0%; Score 47; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10
Db 127 TVSGNLTIR 136

RESULT 3
US-08-791-495-9
; Sequence 9, Application US/08791495
; Patent No. 5811519
; GENERAL INFORMATION:
; APPLICANT: Leth, Bernard
; APPLICANT: Lucas, Sophie
; APPLICANT: De Smet, Charles
; APPLICANT: Godelaine, Daniele
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,495
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Van Amsterdam, John R.
; REGISTRATION NUMBER: 40,212
; REFERENCE/DOCKET NUMBER: L0461/7005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-791-495-9
Query Match      100.0%; Score 47; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10
Db 127 TVSGNLTIR 136

RESULT 4
US-08-937-263B-8
; Sequence 8, Application US/08937263B
; Patent No. 6274145
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tsang; Scanlan, Matthew;
; APPLICANT: Gure, Ali; Old, Lloyd J.; Jager, Elke;
; APPLICANT: Alexander, Knuth; Drifhout, Jan W.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE
; TITLE OF INVENTION: ENCODING CANCER ASSOCIATED ANTIGEN, THE ANTIGEN
; TITLE OF INVENTION: ITSELF, AND USES THEREOF
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: WordPerfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/937,263B  
FILING DATE: September 15, 1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/725,182  
FILING DATE: October 3, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Sinn, Eric, Patent Agent  
REGISTRATION NUMBER: 40,177  
REFERENCE/DOCKET NUMBER: LUD 5466.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 318-3000  
TELEFAX: (212) 752-5958  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 180  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-937-263B-8

Query Match 100.0%; Score 47; DB 3; Length 180;  
Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10  
Db 127 TVSGNLTIR 136

RESULT 5  
US-09-751-798-8  
Sequence 8, Application: US/09751798  
Patent No. 6525177  
GENERAL INFORMATION:  
APPLICANT: Stoeckert, Elisabeth; Jager, Elke;  
APPLICANT: Chen, Yao-tseng; Scanlan, Matthew;  
APPLICANT: Knuth, Alexander; Old, Lloyd J.  
TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer  
TITLE OF INVENTION: Associated Proteins, Uses Thereof,  
TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA  
TITLE OF INVENTION: Binding Peptides Derived Therefrom  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski, L.L.P.  
STREET: 666 Fifth Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: WordPerfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/751,798  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/062,422  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/725,182  
FILING DATE: October 3, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6525177man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5466.3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 318-3168

TELEFAX: (212) 752-5958  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 180  
TYPE: amino acid  
TOPOLOGY: linear  
US-09-751-798-8

Query Match 100.0%; Score 47; DB 4; Length 180;  
Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10  
Db 127 TVSGNLTIR 136

RESULT 6  
US-09-392-714-25  
Sequence 25, Application: US/09392714A  
Patent No. 6686147  
GENERAL INFORMATION:  
APPLICANT: Scanlan, Matthew J.  
APPLICANT: Gure, Ali O.  
APPLICANT: Williamson, Barbara  
APPLICANT: Chen, Yao-Tseng  
APPLICANT: Old, Lloyd J.  
TITLE OF INVENTION: Cancer Associated Antigens and Uses  
TITLE OF INVENTION: Therefor  
FILE REFERENCE: L0461/7062  
CURRENT APPLICATION NUMBER: US/09/392,714A  
CURRENT FILING DATE: 1999-09-09  
EARLIER APPLICATION NUMBER: PCT/US98/14679  
EARLIER FILING DATE: 1998-07-15  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 25  
LENGTH: 180  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-392-714-25

Query Match 100.0%; Score 47; DB 4; Length 180;  
Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10  
Db 127 TVSGNLTIR 136

RESULT 7  
US-09-344-040C-125  
Sequence 125, Application: US/09344040C  
Patent No. 6548064  
GENERAL INFORMATION:  
APPLICANT: Tureci, Ozlem  
APPLICANT: Sahin, Ugur  
APPLICANT: Pfreundschuh, Michael  
APPLICANT: Ramensee, Hans Georg  
APPLICANT: Stevanovic, Stefan  
TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Determining Presence of an SSX Gene, Peptides Derived From Said SSX Gene ;  
TITLE OF INVENTION: Expression of an SSX Gene, and Uses Thereof  
FILE REFERENCE: LUD 5556.1  
CURRENT APPLICATION NUMBER: US/09/344,040C  
CURRENT FILING DATE: 1999-06-25  
PRIOR APPLICATION NUMBER: US 09/105,839  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: US 08/851,130  
PRIOR FILING DATE: 1997-05-05  
NUMBER OF SEQ ID NOS: 132  
SEQ ID NO 125

```
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-344-040C-125

Query Match      89.4%; Score 42; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVSGNLT 9
Db 1 TVSGNLT 9

RESULT 8
US-09-833-039A-125
; Sequence 125, Application US/09833039A
; Patent No. 6673350
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Tumor Associated Peptide and Uses Thereof
; FILE REFERENCE: LUD 5622.1
; CURRENT APPLICATION NUMBER: US/09/833,039A
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US 09/409,455
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/344,040
; PRIOR FILING DATE: 1999-06-26
; PRIOR APPLICATION NUMBER: US 09/105,839
; PRIOR FILING DATE: 1998-06-26
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 125
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-039A-125

Query Match      89.4%; Score 42; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVSGNLT 9
Db 1 TVSGNLT 9

RESULT 9
US-09-344-040C-120
; Sequence 120, Application US/09344040C
; Patent No. 6548064
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; APPLICANT: Ramensee, Hans Georg
; APPLICANT: Stevanovic, Stefan
; TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Determining Expression of an SSX Gene, Peptides Derived From Said SSX Gene and Uses Thereof
; FILE REFERENCE: LUD 5556.1
; CURRENT APPLICATION NUMBER: US/09/344,040C
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,839
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US 08/851,130
; PRIOR FILING DATE: 1997-05-05
; NUMBER OF SEQ ID NOS: 132
; SEQ ID NO 120
; LENGTH: 9
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-344-040C-120

Query Match      80.9%; Score 38; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVSGNLT 8
Db 2 TVSGNLT 9

RESULT 10
US-09-833-039A-120
; Sequence 120, Application US/09833039A
; Patent No. 6673350
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Tumor Associated Peptide and Uses Thereof
; FILE REFERENCE: LUD 5622.1
; CURRENT APPLICATION NUMBER: US/09/833,039A
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US 09/409,455
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/344,040
; PRIOR FILING DATE: 1999-06-26
; PRIOR APPLICATION NUMBER: US 09/105,839
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 120
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-039A-120

Query Match      80.9%; Score 38; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVSGNLT 8
Db 2 TVSGNLT 9

RESULT 11
US-08-791-495-7
; Sequence 7, Application US/08791495
; Patent No. 5811519
; GENERAL INFORMATION:
; APPLICANT: Leth, Bernard
; APPLICANT: Lucas, Sophie
; APPLICANT: De Smet, Charles
; APPLICANT: Godelaine, Daniele
; APPLICANT: Boon-Fallieur, Thierry
; TITLE OF INVENTION: IL-1 TUMOR SPECIFIC GENES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacke, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,495
```

```
/
/
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Van Amsterdam, John R.
/ REGISTRATION NUMBER: 40,212
/ REFERENCE/DOCKET NUMBER: L0461/7005
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-720-3500
/ TELEFAX: 617-720-2441
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 180 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/
US-08-791-495-7

Query Match      80.9%; Score 38; DB 2; Length 180;
Best Local Similarity 80.0%; Pred. No. 5.8;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 TVSGNLTIR 10
Db      127 TVSGNLTIR 136

RESULT 12
US-09-134-001C-5428
; Sequence 5428, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5428
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5428

Query Match      74.5%; Score 35; DB 4; Length 342;
Best Local Similarity 70.0%; Pred. No. 46;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 TVSGNLTIR 10
Db      251 TVGNIMKIR 260

RESULT 13
US-09-540-236-2614
; Sequence 2614, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Briston et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2614
; LENGTH: 149
; TYPE: PRT
; ORGANISM: M.catarrhalis
```

```
US-09-540-236-2614

Query Match      72.3%; Score 34; DB 4; Length 149;
Best Local Similarity 60.0%; Pred. No. 28;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      1 TVSGNLTIR 10
Db      129 TVGNITIK 138

RESULT 14
US-09-732-210-1076
; Sequence 1076, Application US/09732210
; Patent No. 8573361
; GENERAL INFORMATION:
; APPLICANT: Bunkers, Greg J.
; APPLICANT: Liang, Jihong
; APPLICANT: Mittanck, Cindy A.
; APPLICANT: Seale, Jeffrey W.
; APPLICANT: Wu, Yomie S.
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
; FILE REFERENCE: 38-21(15036)B
; CURRENT APPLICATION NUMBER: US/09/732,210
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,513
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 60/169,340
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 1753
; SEQ ID NO 1076
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus
US-09-732-210-1076

Query Match      72.3%; Score 34; DB 4; Length 177;
Best Local Similarity 55.6%; Pred. No. 34;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      1 TVSGNLTIR 9
Db      43 TVGNVITV 51

RESULT 15
US-08-118-270-329
; Sequence 329, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEWMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,270
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
```

ATTORNEY/AGENT INFORMATION:  
NAME: Townseng, Kevin G.  
REGISTRATION NUMBER: 34,033  
REFERENCE/DOCKET NUMBER: MURPHY-2A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 329:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 97 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-118-270-329

Query Match 70.2%; Score 33; DB 1; Length 97;  
Best Local Similarity 66.7%; Pred. No. 27;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TVSGNLT I 9  
Db 17 TIAGNLT VI 25

Search completed: June 22, 2004, 08:54:48  
Job time : 11 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.  
OM protein - protein search, using sw model  
Run on: June 22, 2004, 08:51:28 ; Search time 26 seconds  
(without alignments)  
106.582 Million cell updates/sec

Title: US-09-529-206D-4\_COPY\_127\_136  
Perfect score: 47  
Sequence: 1 TVSGNLTIR 10  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 1163542 seqs, 282313646 residues  
Total number of hits satisfying chosen parameters: 1163542  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	20	15	US-10-164-121A-35
2	47	100.0	20	15	US-10-164-078A-34
3	47	100.0	30	12	US-10-296-734-1414
4	47	100.0	54	16	US-10-447-161-140
5	47	100.0	180	9	US-09-751-798-8
6	47	100.0	180	10	US-09-849-602-30
7	47	100.0	180	12	US-10-296-734-832
8	47	100.0	180	13	US-10-023-182-8
9	47	100.0	180	14	US-10-207-655-71
10	47	100.0	180	15	US-10-026-066-3
11	47	100.0	180	15	US-10-117-937-74
12	47	100.0	180	15	US-10-295-027-386
13	47	100.0	180	16	US-10-188-832-139
14	47	100.0	397	9	US-09-821-893-27
15	47	100.0	3541	12	US-10-296-734-1454

16	42	89.4	9	10	US-09-833-039-125	Sequence 125, App
17	42	89.4	9	14	US-10-177-277-125	Sequence 125, App
18	42	89.4	9	15	US-10-117-937-185	Sequence 185, App
19	42	89.4	10	15	US-10-117-937-186	Sequence 186, App
20	38	80.9	9	10	US-09-833-039-120	Sequence 120, App
21	38	80.9	9	14	US-10-177-277-120	Sequence 120, App
22	38	80.9	30	12	US-10-296-734-1446	Sequence 1446, Ap
23	38	80.9	180	12	US-10-296-734-834	Sequence 834, App
24	38	80.9	180	14	US-10-146-473-69	Sequence 69, Appl
25	38	80.9	180	15	US-10-117-937-75	Sequence 75, Appl
26	37	78.7	330	12	US-10-282-122A-61146	Sequence 61146, A
27	35	74.5	317	12	US-10-092-900A-196	Sequence 196, App
28	35	74.5	334	12	US-10-282-122A-71904	Sequence 71904, A
29	35	74.5	335	9	US-09-815-242-12730	Sequence 12730, A
30	35	74.5	335	12	US-10-282-122A-70259	Sequence 70259, A
31	35	74.5	336	12	US-10-282-122A-70661	Sequence 70661, A
32	35	74.5	742	13	US-10-108-605-203	Sequence 203, App
33	34	72.3	49	12	US-10-424-599-174030	Sequence 174030,
34	34	72.3	132	15	US-10-074-978A-399	Sequence 399, App
35	34	72.3	177	13	US-10-013-379-6	Sequence 6, Appli
36	34	72.3	299	14	US-10-060-795B-6	Sequence 242, App
37	34	72.3	585	12	US-10-219-535-242	Sequence 242, App
38	34	72.3	585	12	US-10-232-230-242	Sequence 242, App
39	34	72.3	585	12	US-10-232-224-242	Sequence 242, App
40	34	72.3	585	14	US-10-227-884-242	Sequence 242, App
41	34	72.3	585	14	US-10-230-163-242	Sequence 242, App
42	34	72.3	585	14	US-10-230-338-242	Sequence 242, App
43	34	72.3	585	14	US-10-218-631-242	Sequence 242, App
44	34	72.3	585	14	US-10-230-414-242	Sequence 242, App
45	34	72.3	585	14	US-10-216-159A-242	Sequence 242, App

ALIGNMENTS

RESULT 1  
US-10-164-121A-35  
; Sequence 35, Application US/10164121A  
; Publication No. US20030228308A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Yi  
; APPLICANT: Boon, Thierry  
; APPLICANT: Van der Bruggen, Pierre  
; APPLICANT: Traversari, Catra  
; TITLE OF INVENTION: Isolated Peptides Which Bind to HLA-Cw6 Molecules And Uses Thereof  
; FILE REFERENCE: LUD-5771  
; CURRENT APPLICATION NUMBER: US/10/164,121A  
; CURRENT FILING DATE: 2002-08-26  
; NUMBER OF SEQ ID NOS: 36  
; SEQ ID NO 35  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: H. sapiens  
; FEATURE:  
US-10-164-121A-35

Query Match 100.0%; Score 47; DB 15; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.033; 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10  
|||||  
Db 9 TVSGNLTIR 18

RESULT 2  
US-10-164-078A-34  
; Sequence 34, Application US/10164078A  
; Publication No. US20030228325A1  
; GENERAL INFORMATION:  
; APPLICANT: Bileborough, Janine  
; APPLICANT: Schultz, Erwin  
; APPLICANT: Panichelli, Christophe

; APPLICANT: Van der Bruggen, Pierre  
; APPLICANT: Boon, Thierry  
; TITLE OF INVENTION: Isolated Peptides Which Bind to HLA-B18 Molecules And Uses Thereof  
; FILE REFERENCE: LUD-5756  
; CURRENT APPLICATION NUMBER: US/10/164,078A  
; CURRENT FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 35  
; SEQ ID NO 34  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: H. sapiens  
; FEATURE:  
US-10-164-078A-34

Query Match 100.0%; Score 47; DB 15; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10  
Db 9 TVSGNLTIR 18

RESULT 3  
US-10-296-734-1414  
; Sequence 1414, Application US/10296734  
; Publication No. US20040054137A1  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Scott A  
; APPLICANT: Ramshaw, Ian A  
; TITLE OF INVENTION: Synthetic molecules and uses therefor  
; FILE REFERENCE: Savine  
; CURRENT APPLICATION NUMBER: US/10/296,734  
; CURRENT FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: AU P07761/00  
; PRIOR FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 1507  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1414  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: NYSOLA segment 9  
US-10-296-734-1414

Query Match 100.0%; Score 47; DB 12; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.051;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10  
Db 9 TVSGNLTIR 18

RESULT 4  
US-10-447-161-140  
; Sequence 140, Application US/10447161  
; Publication No. US2004002314A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Rong-fu  
; TITLE OF INVENTION: Mutant Fibronectin and Tumor Metastasis  
; FILE REFERENCE: HQ-P02484US1  
; CURRENT APPLICATION NUMBER: US/10/447,161  
; CURRENT FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/383,530  
; PRIOR FILING DATE: 2002-05-28  
; NUMBER OF SEQ ID NOS: 148  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 140  
; LENGTH: 54  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

; FEATURE:  
; OTHER INFORMATION: Synthetic Peptide  
US-10-447-161-140

Query Match 100.0%; Score 47; DB 16; Length 54;  
Best Local Similarity 100.0%; Pred. No. 0.097;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10  
Db 25 TVSGNLTIR 34

RESULT 5  
US-09-751-798-8  
; Sequence 8, Application US/09751798  
; Patent No. US20020010321A1  
; GENERAL INFORMATION:  
; APPLICANT: Stockert, Elisabeth; Jager, Elke;  
; APPLICANT: Chen, Yao-tseng; Scanlan, Matthew;  
; APPLICANT: Knuth, Alexander; Old, Lloyd J.  
; TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer  
; TITLE OF INVENTION: Associated Proteins, Uses Thereof,  
; TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA  
; TITLE OF INVENTION: Binding Peptides Derived Therefrom  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fulbright & Jaworski, L.L.P.  
; STREET: 666 Fifth Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/751,798  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/062,422  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/725,182  
; FILING DATE: October 3, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. US20020010321A1man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: LUD 5466.3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 318-3168  
; TELEFAX: (212) 752-5958  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 180  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-09-751-798-8

Query Match 100.0%; Score 47; DB 9; Length 180;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10  
Db 127 TVSGNLTIR 136

RESULT 6  
US-09-849-602-30  
; Sequence 30, Application US/09849602

```
; Publication No. US20030165834A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Old, Lloyd J.
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Chen, Yao-Tseng
; TITLE OF INVENTION: Colon Cancer Antigen Panel
; FILE REFERENCE: L0461/7105(JRV)
; CURRENT APPLICATION NUMBER: US/09/849,602
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-602-30

Query Match      100.0%; Score 47; DB 10; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TVSGNLTIR 10
DB      127 TVSGNLTIR 136

RESULT 7
US-10-296-734-832
; Sequence 832, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 832
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: NYNSOLA consensus polypeptide
US-10-296-734-832

Query Match      100.0%; Score 47; DB 12; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TVSGNLTIR 10
DB      127 TVSGNLTIR 136

RESULT 8
US-10-023-182-8
; Sequence 8, Application US/10023182
; Publication No. US20020164665A1
; GENERAL INFORMATION:
; APPLICANT: Stockert, Elisabeth; Jager, Elke;
; Chen, Yao-tseng; Scanlan, Matthew;
; Knuth, Alexander; Old, Lloyd J.
; TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer
; Associated Proteins, Uses Thereof,
; Truncated Forms of NY-ESO-1, and HLA
; Binding Peptides Derived Therefrom
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/023,182
; FILING DATE: 17-Dec-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/751,798
; FILING DATE: December 29, 2000
; APPLICATION NUMBER: 09/062,422
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/725,182
; FILING DATE: October 3, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. US20020164665Alman D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5466.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3168
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-023-182-8

Query Match      100.0%; Score 47; DB 13; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TVSGNLTIR 10
DB      127 TVSGNLTIR 136

RESULT 9
US-10-207-655-71
; Sequence 71, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A., Martha S.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390059.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 71
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-71

Query Match      100.0%; Score 47; DB 14; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TVSGNLTIR 10
DB      127 TVSGNLTIR 136
```

## RESULT 10

US-10-026-066-3  
 ; Sequence 3, Application US/10026066  
 ; Publication No. US20030215425A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Simard, John J. L.  
 ; APPLICANT: Diamond, David C.  
 ; TITLE OF INVENTION: EPITOPE SYNCHRONIZATION IN ANTIGEN  
 ; FILE REFERENCE: CTILIM.21CPIIC  
 ; CURRENT FILING DATE: 2001-12-07  
 ; CURRENT APPLICATION NUMBER: US/10/026,066  
 ; PRIOR FILING DATE: 2001-04-28  
 ; PRIOR APPLICATION NUMBER: 09/561,074  
 ; PRIOR FILING DATE: 2000-04-28  
 ; PRIOR APPLICATION NUMBER: 09/560,465  
 ; PRIOR FILING DATE: 2000-04-28  
 ; PRIOR APPLICATION NUMBER: 09/561,572  
 ; PRIOR FILING DATE: 2000-04-28  
 ; PRIOR APPLICATION NUMBER: 09/561,571  
 ; PRIOR FILING DATE: 2000-04-28  
 ; PRIOR APPLICATION NUMBER: PCT/US01/13806  
 ; PRIOR FILING DATE: 2001-04-27  
 ; NUMBER OF SEQ ID NOS: 89  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 180  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-026-066-3

Query Match 100.0%; Score 47; DB 15; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 0.36;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10

Db 127 TVSGNLTIR 136

## RESULT 11

US-10-117-937-74  
 ; Sequence 74, Application US/10117937  
 ; Publication No. US20030220239A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CTL IMMUNO THERAPIES CORP.  
 ; APPLICANT: SIMARD, John, J.L.  
 ; APPLICANT: DIAMOND, David, C.  
 ; APPLICANT: LIU, Liping  
 ; APPLICANT: XIE, Zhidong  
 ; TITLE OF INVENTION: EPITOPE SEQUENCES  
 ; FILE REFERENCE: CTILIM.022A  
 ; CURRENT FILING DATE: 2002-04-04  
 ; CURRENT APPLICATION NUMBER: US/10/117,937  
 ; PRIOR FILING DATE: 2002-04-04  
 ; PRIOR APPLICATION NUMBER: US 60/282,211  
 ; PRIOR FILING DATE: 2001-04-06  
 ; PRIOR APPLICATION NUMBER: US 60/337,017  
 ; PRIOR FILING DATE: 2001-11-07  
 ; PRIOR APPLICATION NUMBER: US 60/363,210  
 ; PRIOR FILING DATE: 2002-03-07  
 ; NUMBER OF SEQ ID NOS: 602  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 74  
 ; LENGTH: 180  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-117-937-74

Query Match 100.0%; Score 47; DB 15; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 0.36;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10

Db 127 TVSGNLTIR 136

## Db

127 TVSGNLTIR 136

## RESULT 12

US-10-295-027-386  
 ; Sequence 386, Application US/10295027  
 ; Publication No. US20030212350A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Afar, Daniel  
 ; APPLICANT: Aziz, Natasha  
 ; APPLICANT: Ginsberg, Wendy M.  
 ; APPLICANT: Gish, Kurt C.  
 ; APPLICANT: Glynn, Richard  
 ; APPLICANT: Hevezi, Peter A.  
 ; APPLICANT: Mack, David H.  
 ; APPLICANT: Murray, Richard  
 ; APPLICANT: Watson, Susan R.  
 ; APPLICANT: Eos Biotechnology, Inc.  
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
 ; FILE REFERENCE: 018501-012500US  
 ; CURRENT FILING DATE: 2002-11-13  
 ; CURRENT APPLICATION NUMBER: US/10/295,027  
 ; PRIOR FILING DATE: 2000-09-15  
 ; PRIOR APPLICATION NUMBER: US 60/350,666  
 ; PRIOR FILING DATE: 2001-11-13  
 ; PRIOR APPLICATION NUMBER: US 60/335,394  
 ; PRIOR FILING DATE: 2001-11-15  
 ; PRIOR APPLICATION NUMBER: US 60/332,464  
 ; PRIOR FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: US 60/334,393  
 ; PRIOR FILING DATE: 2001-11-29  
 ; PRIOR APPLICATION NUMBER: US 60/340,376  
 ; PRIOR FILING DATE: 2001-12-14  
 ; PRIOR APPLICATION NUMBER: US 60/347,211  
 ; PRIOR FILING DATE: 2002-01-08  
 ; PRIOR APPLICATION NUMBER: US 60/347,349  
 ; PRIOR FILING DATE: 2002-01-10  
 ; PRIOR APPLICATION NUMBER: US 60/355,250  
 ; PRIOR FILING DATE: 2002-02-08  
 ; PRIOR APPLICATION NUMBER: US 60/356,714  
 ; PRIOR FILING DATE: 2002-02-13  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 1386  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 386  
 ; LENGTH: 180  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-295-027-386

Query Match 100.0%; Score 47; DB 15; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 0.36;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10

Db 127 TVSGNLTIR 136

## RESULT 13

US-10-188-832-139  
 ; Sequence 139, Application US/10188832  
 ; Publication No. US20040076955A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mack, David H.  
 ; APPLICANT: Aziz, Natasha  
 ; APPLICANT: Eos Biotechnology, Inc.  
 ; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions  
 ; FILE REFERENCE: and Methods of Screening for Modulators of Bladder  
 ; TITLE OF INVENTION: Cancer  
 ; FILE REFERENCE: 018501-002330US

; CURRENT APPLICATION NUMBER: US/10/188,832  
; CURRENT FILING DATE: 2002-11-22  
; PRIOR APPLICATION NUMBER: US 60/302,814  
; PRIOR FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: US 60/310,099  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: US 60/343,705  
; PRIOR FILING DATE: 2001-11-08  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/372,246  
; PRIOR FILING DATE: 2002-04-12  
; NUMBER OF SEQ ID NOS: 207  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 139  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-188-832-139

Query Match 100.0%; Score 47; DB 16; Length 180;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 TVSGNLTIR 10  
Db 127 TVSGNLTIR 136

RESULT 14  
US-03-821-883-27  
; Sequence 27, Application US/09621883  
; Patent No. US20020061310A1  
; GENERAL INFORMATION:  
; APPLICANT: Laus, Reiner  
; APPLICANT: Vidovic, Damir  
; APPLICANT: Graddis, Thomas  
; TITLE OF INVENTION: Cell-Based Immunotherapy  
; TITLE OF INVENTION: Compositions and Methods for Dendritic  
; FILE REFERENCE: 7636-0022.30  
; CURRENT APPLICATION NUMBER: US/09/821,883  
; CURRENT FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: US 60/193,504  
; PRIOR FILING DATE: 2000-03-30  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 27  
; LENGTH: 397  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: NY-ESO-1C tumor antigen  
US-09-821-883-27

Query Match 100.0%; Score 47; DB 9; Length 397;  
Best Local Similarity 100.0%; Pred. No. 0.86;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 TVSGNLTIR 10  
Db 127 TVSGNLTIR 136

RESULT 15  
US-10-296-734-1454  
; Sequence 1454, Application US/10296734  
; Publication No. US20040054137A1  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Scott A  
; APPLICANT: Ramshaw, Ian A  
; TITLE OF INVENTION: Synthetic molecules and uses therefor  
; FILE REFERENCE: Savine  
; CURRENT APPLICATION NUMBER: US/10/296,734

; CURRENT FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: AU PQ7761/00  
; PRIOR FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 1507  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1454  
; LENGTH: 3541  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Melanoma cancer specific savine  
US-10-296-734-1454  
  
Query Match 100.0%; Score 47; DB 12; Length 3541;  
Best Local Similarity 100.0%; Pred. No. 9.5;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 TVSGNLTIR 10  
Db 2139 TVSGNLTIR 2148  
  
Search completed: June 22, 2004, 08:56:19  
Job time : 27 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2004, 08:48:03 ; Search time 9 Seconds  
(without alignments)  
106.879 Million cell updates/sec

Title: US-09-529-206D-4\_COPY\_127\_136

Perfect score: 47

Sequence: 1 TVSGNLTITV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	76.6	245	2 T13099	major tail protein
2	36	76.6	323	2 AD2753	lipic acid synthase
3	36	76.6	323	2 B97534	lipic acid synthase
4	35	74.5	335	2 A89837	hypothetical protein
5	34	72.3	177	1 R850F	ribosomal protein
6	34	72.3	244	2 T41234	translation initiation
7	34	72.3	311	2 AB3218	2-hydroxyacid dehy
8	34	72.3	314	2 G86835	hypothetical protein
9	34	72.3	1266	2 A95989	hypothetical protein
10	34	72.3	1266	2 F91143	hypothetical protein
11	33	70.2	144	1 T8PPI	tail fiber protein
12	33	70.2	144	2 S18683	Gene R protein - E
13	33	70.2	411	2 G95241	MATE efflux family
14	33	70.2	426	2 B98106	hypothetical protein
15	33	70.2	428	2 F85253	hypothetical protein
16	33	70.2	436	2 A81977	probable amino-acid
17	33	70.2	436	2 B81033	N-acetylglutamate
18	33	70.2	449	2 S02011	serotonin receptor
19	33	70.2	471	2 A34863	serotonin receptor
20	33	70.2	471	2 S40689	5-hydroxytryptamin
21	33	70.2	471	2 S11280	serotonin receptor
22	33	70.2	471	2 A43956	serotonin receptor
23	33	70.2	876	2 E96674	hypothetical protein
24	33	70.2	918	2 T02759	hypothetical protein
25	33	70.2	1896	2 F71405	probable TMV resist
26	32	68.1	102	2 C6898	hypothetical protein
27	32	68.1	121	2 B81972	probable integral
28	32	68.1	135	2 D81029	conserved hypothet
29	32	68.1	187	2 I37105	5-HT2c receptor -

30 32 68.1 234 2 H75390  
31 32 68.1 245 2 S57550  
32 32 68.1 289 2 C87158  
33 32 68.1 316 2 S25234  
34 32 68.1 344 2 A69325  
35 32 68.1 347 2 G95146  
36 32 68.1 347 2 E98014  
37 32 68.1 376 2 C97059  
38 32 68.1 418 2 E70401  
39 32 68.1 451 2 S65162  
40 32 68.1 455 2 HMXRS3  
41 32 68.1 458 2 JS0616  
42 32 68.1 459 2 A43951  
43 32 68.1 460 2 A32605  
44 32 68.1 479 2 S23562  
45 32 68.1 491 2 T30590

## ALIGNMENTS

### RESULT 1

T13099  
major tail protein V - phage N15

N:Alternate names: protein gp13

C:Species: phage N15

C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 13-Aug-1999

C/Accession: T13099

E:Hendrix, R.W.; Ravin, V.K.; Casjens, S.R.; Ford, M.E.; Ravin, N.V.; Smirnov, I.K.

submitted to the EMBL Data Library, May 1998

A:Reference number: Z17603

A:Accession: T13099

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-245 <H2N>

A/Cross-references: EMBL:AF064539; NID:G3192683; PID:G3192697; FIDN:AAC19050.1

C:Genetics:

A:Note: Gene 13

C:Superfamily: phage lambda major tail protein V

Query Match 76.6%; Score 36; DB 2; Length 245;  
Best Local Similarity 77.8%; Pred. No. 13;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TVSGNLTITV 9

Db 209 TVSGNLTITV 217

### RESULT 2

AD2753

lipic acid synthetase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002

C/Accession: AD2753

R:Wood, D.W.; Secubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A/Accession: AD2753

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-323 <KUR>

A/Cross-references: GB:AE008688; PIDN:AL42442.1; PID:G17739856; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: lipA

A:Map position: circular chromosome

C:Superfamily: lipic acid synthase

Query Match 76.6%; Score 36; DB 2; Length 323;  
 Best Local Similarity 70.0%; Pred. No. 18;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10  
 |||||  
 Db 189 TVFGNYLTJR 198

RESULT 3  
 B97534  
 A:Title: lipoic acid synthetase (lip-syn) (lipoate synthase) [imported] - Agrobacterium tumefaciens  
 C:Species: Agrobacterium tumefaciens  
 C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002  
 C:Accession: B97534  
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
 A:Reference number: A97359; MUID:21608551; PMID:11743194  
 A:Accession: B97534  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-323 <KUR>  
 A:Cross-references: GB:AE007869; PIDN:AAK87227.1; PID:gl5156511; GSPDB:GN00169  
 C:Genetics:  
 A:Gene: AGR\_C\_2646  
 A:Map position: circular chromosome  
 C:Superfamily: lipoic acid synthase

Query Match 76.6%; Score 36; DB 2; Length 323;  
 Best Local Similarity 70.0%; Pred. No. 18;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10  
 |||||  
 Db 189 TVFGNYLTJR 198

RESULT 4  
 A89837  
 A:Title: hypothetical protein SA0619 [imported] - Staphylococcus aureus (strain N315)  
 C:Species: Staphylococcus aureus  
 C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 01-Mar-2002  
 C:Accession: A89837  
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaico, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.; Lancet 357, 1225-1240, 2001  
 A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
 A:Reference number: A89758; MUID:21311952; PMID:11418146  
 A:Accession: A89837  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-335 <KUR>  
 A:Cross-references: GB:BA000018; PID:gl3700555; PIDN:BAB41852.1; GSPDB:GN00149  
 A:Experimental source: strain N315  
 C:Genetics:  
 A:Gene: SA0619  
 C:Superfamily: probable sodium-dependent phosphate transporter MTH1885

Query Match 74.5%; Score 35; DB 2; Length 335;  
 Best Local Similarity 70.0%; Pred. No. 29;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10  
 |||||  
 Db 245 TVGYNMKIR 254

RESULT 5  
 R5BS0F

ribosomal protein L6 - Bacillus stearothermophilus  
 N:Alternate names: ribosomal protein BL10  
 C:Species: Bacillus stearothermophilus  
 C>Date: 02-Apr-1982 #sequence\_revision 02-Apr-1982 #text\_change 12-Jun-2003  
 C:Accession: A02766; B39085; S59061  
 R:Kimura, M.; Rawlings, N.; Appelt, K.; FEBS Lett. 136, 58-64, 1981  
 A:Title: The amino acid sequence of protein BL10 from the 50S subunit of the Bacillus s  
 A:Reference number: A02766  
 A:Accession: A02766  
 A:Molecule type: protein  
 A:Residues: 1-177 <KIM>  
 R:Ramakrishnan, V.; Garchman, S.E.; J. Biol. Chem. 266, 880-885, 1991  
 A:Title: Cloning, sequencing, and overexpression of genes for ribosomal proteins from E  
 A:Reference number: A39085; MUID:91093287; PMID:1985969  
 A:Accession: B39085  
 A:Molecule type: DNA  
 A:Residues: 8-170 <RAM>  
 A:Cross-references: GB:M57622  
 R:Urbala, H.; Kluft, V.; Bischof, O.; Mueller, E.C.; Wittmann-Liebold, B.; EMBO J. 14, 4578-4588, 1995  
 A:Title: Protein-rRNA binding features and their structural and functional implications  
 A:Reference number: S59051; MUID:96003638; PMID:7556101  
 A:Accession: S59061  
 A:Molecule type: protein  
 A:Residues: 149-163 <URL>  
 C:Superfamily: ribosomal protein L6/L9  
 C:Keywords: protein biosynthesis; ribosome

Query Match 72.3%; Score 34; DB 1; Length 177;  
 Best Local Similarity 55.6%; Pred. No. 24;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 9  
 |||||  
 Db 43 TVEGNVITV 51

RESULT 6  
 T41234  
 A:Title: translation initiation factor eIF-6 [imported] - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 15-Sep-2000  
 C:Accession: T41234  
 R:Lyne, M.; Harris, D.E.; Murphey, L.D.; Rajandream, M.A.; Barrell, B.G. submitted to the EMBL Data Library, January 1999  
 A:Reference number: Z21979  
 A:Accession: T41234  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-244 <LYN>  
 A:Cross-references: EMBL:AL035075; PIDN:CAA22640.1; GSPDB:GN00068; SPDB:SPCC1919.09  
 A:Experimental source: strain 972h-; cosmid c1919  
 C:Genetics:  
 A:Gene: SPDB:SPCC1919.09  
 A:Map position: 3  
 A:Introns: 12/2; 36/2  
 C:Superfamily: conserved hypothetical protein YPR016c

Query Match 72.3%; Score 34; DB 2; Length 244;  
 Best Local Similarity 75.0%; Pred. No. 33;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 8  
 |||||  
 Db 141 TVAGNVLT 148

RESULT 7  
 AB3218  
 A:Title: 2-hydroxyacid dehydrogenase Atu5474 [imported] - Agrobacterium tumefaciens (strain C59,  
 C:Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
C;Accession: AB3218  
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A;Reference number: AB2577; MUID:21608550; PMID:11743193  
A;Accession: AB3218  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-311 <KUR>  
A;Cross-references: GB:AB08687; PIDN:AA46160.1; PID:g17743930; GSPDB:GN00188  
A;Experimental source: strain C58 (Dupont)  
C;Genetics:  
A;Gene: Atu5474  
A;Genome: plasmid  
C;Superfamily: phosphoglycerate dehydrogenase  
  
Query Match 72.3%; Score 34; DB 2; Length 311;  
Best Local Similarity 77.8%; Pred. No. 43;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 2 VSGNLTIR 10  
|:|||||  
Db 288 VAGNLTIR 296  
  
RESULT 8  
GB835  
A;Title: Lactococcus lactis subsp. lactis (strain IL1403)  
C;Species: Lactococcus lactis subsp. lactis  
C;Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C;Accession: G86835  
R;Bolotin, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrli  
Genome Res. 11, 731-753, 2001  
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
A;Reference number: AB6625; MUID:21235186; PMID:11337471  
A;Accession: G86835  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-314 <STO>  
A;Cross-references: GB:AB005176; PID:g12724702; PIDN:AAK05785.1; GSPDB:GN00146  
A;Experimental source: strain IL1403  
C;Genetics:  
A;Gene: yreB  
  
Query Match 72.3%; Score 34; DB 2; Length 314;  
Best Local Similarity 75.0%; Pred. No. 44;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 2 VSGNLTIR 9  
|:|||||  
Db 93 ISGNLTIL 100  
  
RESULT 9  
A85989  
A;Title: Hypothetical protein yreB [imported] - Escherichia coli (strain O157:H7, substrain EDL9  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 17-May-2002  
C;Accession: A85989  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: A85989  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1266 <STO>

A;Cross-references: GB:AE005174; NID:g12517869; PIDN:AAG58373.1; GSPDB:GN00145; UWGP:241  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: Z4604  
C;Superfamily: Escherichia coli hypothetical protein yndp  
  
Query Match 72.3%; Score 34; DB 2; Length 1266;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 TVSGNLTIR 9  
|:|||||  
Db 1017 TISGNLTIL 1025  
  
RESULT 10  
F91143  
A;Title: Hypothetical protein ECs4118 [imported] - Escherichia coli (strain O157:H7, substrain R  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 17-May-2002  
C;Accession: F91143  
R;Hayaishi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen  
A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Accession: F91143  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1266 <HAY>  
A;Cross-references: GB:BA000007; PIDN:BA37541.1; PID:g13363591; GSPDB:GN00154  
A;Experimental source: strain O157:H7, substrain RMD 0509952  
C;Genetics:  
A;Gene: ECs4118  
C;Superfamily: Escherichia coli hypothetical protein yndp  
  
Query Match 72.3%; Score 34; DB 2; Length 1266;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 TVSGNLTIR 9  
|:|||||  
Db 1017 TISGNLTIL 1025  
  
RESULT 11  
TPBPPI  
A;Title: Tail fiber protein R - phage P1  
C;Species: phage P1  
C;Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 23-Jul-1999  
C;Accession: JS0460  
R;Guidolin, A.; Zingg, J.M.; Arber, W.  
Gene 76, 239-243, 1989  
A;Title: Organization of the bacteriophage P1 tail-fibre operon.  
A;Reference number: PS0109; MUID:89326122; PMID:2526777  
A;Accession: JS0460  
A;Molecule type: DNA  
A;Residues: 1-144 <GUI>  
A;Cross-references: GB:M25470; NID:g341349; PIDN:AAAS8777.1; PID:g538421  
C;Genetics:  
A;Gene: R  
C;Superfamily: phage P1 tail fiber protein R  
C;Keywords: tail fiber  
  
Query Match 70.2%; Score 33; DB 1; Length 144;  
Best Local Similarity 60.0%; Pred. No. 30;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 TVSGNLTIR 10  
|:|||||  
Db 79 TVSGLTIR 88

RESULT 12

S18683

gene R protein - Escherichia coli plasmid p15B

C:Species: Escherichia coli

C:Date: 03-May-1994 #sequence\_revision 20-Feb-1995 #text\_change 20-Sep-1999

C:Accession: S18683

R:Sandmeier, H.; Iida, S.; Huebner, P.; Hiestand-Nauer, R.; Arber, W.

Nucleic Acids Res. 19, 5831-5838, 1991

A:Title: Gene organization in the multiple DNA inversion region Min of plasmid p15B of E

A:Reference number: S18680; MUID:92051368; PMID:1945872

A:Accession: S18683

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-144 <SAN>

A:Cross-references: EMBL:X62121; NID:g42224; PIDN:CAA44046.1; PID:g42228

C:Genetics:

A:Gene: R

A:Genome: plasmid

C:Superfamily: phage P1 tail fiber protein R

Query Match 70.2%; Score 33; DB 2; Length 144;

Best Local Similarity 60.0%; Pred. No. 30;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10

DB 79 TVSGGIIIVR 88

RESULT 13

G95241

-MATE efflux family protein [imported] - Streptococcus pneumoniae (strain TIGR4)

C:Species: Streptococcus pneumoniae

C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 03-Aug-2001

C:Accession: G95241

R:Tetzelin, H.; Nelson, K.B.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf,

nson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: G95241

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-411 <KUR>

A:Cross-references: GB:AE005672; PIDN:AAK76128.1; PID:g14973575; GSPDB:GN00164; TIGR:SP4

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP2065

Query Match 70.2%; Score 33; DB 2; Length 411;

Best Local Similarity 66.7%; Pred. No. 92;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 9

DB 30 SVAGNIITI 38

RESULT 14

B98106

hypothetical protein MATE transporter [imported] - Streptococcus pneumoniae (strain R6)

C:Species: Streptococcus pneumoniae

C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 22-Oct-2001

C:Accession: B98106

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: B98106

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-426 <KUR>

A:Cross-references: GB:AE007317; PIDN:AAL00679.1; PID:g15459568; GSPDB:GN00174

C:Genetics:

A:Gene: MATE transporter

Query Match 70.2%; Score 33; DB 2; Length 426;

Best Local Similarity 66.7%; Pred. No. 96;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 9

DB 45 SVAGNIITI 53

RESULT 15

P85253

hypothetical protein AT4g22170 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 02-Mar-2001

C:Accession: P85253

R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spr

ing

Nature 402, 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: A85001; MUID:20083488; PMID:10617198

A:Accession: P85253

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-428 <STO>

A:Cross-references: GB:NC\_001268; NID:g7269062; PIDN:CAB79172.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4g22170

A:Map position: 4

C:Superfamily: Arabidopsis thaliana hypothetical protein T12H17.50

Query Match 70.2%; Score 33; DB 2; Length 428;

Best Local Similarity 60.0%; Pred. No. 96;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10

DB 316 TVTGNILRVQ 325

Search completed: June 22, 2004, 08:54:04

Job time : 11 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: June 22, 2004, 08:41:07 ; Search time 7 Seconds  
(without alignments)  
74,386 Million cell updates/sec

Title: US-09-529-206d-4\_COPY\_127\_136

Perfect score: 47

Sequence: 1 TVSGNLTIR 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	180	1 CTG1_HUMAN	P78358 homo sapien
2	38	80.9	322	1 LIPA_RHET	Q05941 rhizobium e
3	36	76.6	323	1 LIPA_RGRTS	Q8ufg1 agrobacteri
4	34	72.3	131	1 PARA_PIG	Q97788 sus scrofa
5	34	72.3	177	1 RL6_BACST	P02391 bacillus st
6	34	72.3	244	1 IF6_SCHPO	Q94476 schizosacch
7	34	72.3	314	1 OR23_MOUSE	Q8vg09 mus musculu
8	34	72.3	314	1 OR25_MOUSE	Q8vg07 mus musculu
9	34	72.3	314	1 OR28_MOUSE	Q8vg04 mus musculu
10	34	72.3	314	1 OR32_MOUSE	Q8vfd2 mus musculu
11	34	72.3	314	1 OR33_MOUSE	Q8vfd1 mus musculu
12	34	72.3	314	1 OR36_MOUSE	Q8vfe6 mus musculu
13	34	72.3	314	1 OR39_MOUSE	Q8vew6 mus musculu
14	34	72.3	314	1 OR40_MOUSE	Q8vew5 mus musculu
15	34	72.3	316	1 OR22_MOUSE	Q8vg13 mus musculu
16	34	72.3	321	1 OR31_MOUSE	Q8vfd3 mus musculu
17	34	72.3	323	1 OR29_MOUSE	Q8vg03 mus musculu
18	34	72.3	330	1 OR41_MOUSE	Q8vew2 mus musculu
19	34	72.3	330	1 OR42_MOUSE	Q8vfi2 mus musculu
20	33	70.2	144	1 VGR_BPP1	P22946 bacterioph
21	33	70.2	145	1 PHEE_CLOPE	Q8xkil clostridium
22	33	70.2	313	1 O4E2_HUMAN	Q8gcq2 homo sapien
23	33	70.2	436	1 ARGA_NEIMA	Q91w21 neisseria m
24	33	70.2	436	1 ARGA_NEIMB	Q91xu9 neisseria m
25	33	70.2	470	1 SH2A_PIG	P50129 sus scrofa
26	33	70.2	471	1 SH2A_CRIGR	P18599 cricetus
27	33	70.2	471	1 SH2A_HUMAN	P28223 homo sapien
28	33	70.2	471	1 SH2A_NACMU	P50128 macaca mula
29	33	70.2	471	1 SH2A_MOUSE	P35363 mus musculu
30	33	70.2	471	1 SH2A_RAT	P14842 rattus norv
31	32	68.1	104	1 CVAY_VIBPA	Q87k1 vibrio para
32	32	68.1	131	1 PARA_SPERTR	Q99p60 spermophilu
33	32	68.1	209	1 HIS2_PYRFR	P58853 pyrococcus

ALIGNMENTS

RESULT 1

ID	CTG1_HUMAN	STANDARD;	PRT;	180 AA.
AC	P78358;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Cancer/testis antigen 1 (Autoimmunogenic cancer/testis antigen NY-ESO-1).			
GN	CTAG1 OR CTAG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97203161; PubMed=9050879;			
RA	Chen Y.-T., Scanlan M.J., Sahin U., Tuerci O., Gure A.O., Tsang S.,			
RA	Williamson B., Stockert E., Pfrendschuh M., Old L.J.;			
RT	"A testicular antigen aberrantly expressed in human cancers detected			
RT	by autologous antibody screening.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 94:1914-1918 (1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Melanoma;			
RX	MEDLINE=98289662; PubMed=9626360;			
RA	Lethe B., Lucas S., Michaux L., de Smet C., Godelaine D., Serrano A.,			
RA	de Plaen E., Boon T.;			
RT	"LAGB-1, a new gene with tumor specificity.";			
RL	Int. J. Cancer 76:903-908 (1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98430682; PubMed=9759882;			
RA	Wang R.-P., Johnston S.L., Zeng G., Topalian S.L.,			
RA	Schwartzentruber D.J., Rosenberg S.A.;			
RT	"A breast and melanoma-shared tumor antigen: T cell responses to			
RT	antigenic peptides translated from different open reading frames.";			
RL	J. Immunol. 161:3596-3606 (1998).			
CC	-I- TISSUE SPECIFICITY: Expressed in testis and ovary and in a wide			
CC	variety of cancers. Detected in uterine myometrium.			
CC	-I- SIMILARITY: Belongs to the CTAG family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; U87459; AAB49693.1; -			
DR	EMBL; AJ003149; CAA05908.1; -			
DR	EMBL; AF038567; AAD05202.1; -			
DR	Genem; HGNC:2491; CTAG1.			
DR	MIM; 300156; -			
KW	Transmembrane; Antigen.			

```
FT DOMAIN 5 82 GLY-RICH.
FT TRANSMEM 156 172 POTENTIAL.
SQ SEQUENCE 180 AA; 17992 MW; B122C5C2C8BE1569 CRC64;

Query Match 100.0%; Score 47; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10
DB 127 TVSGNLTIR 136

RESULT 2
LIPA RHET STANDARD; PRT; 322 AA.
AC O05941;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lipoic acid synthetase (Lip-syn) (Lipoate synthase).
GN LIPA.
OS Rhizobium etli.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=29449;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CE3;
RX MEDLINE=97286510; PubMed=9141657;
RA Tate R., Riccio A., Iaccarino M., Patriarca E.J.;
RT "Cloning and transcriptional analysis of the lipA (lipoic acid
RL PFMS Microbiol. Lett. 149:165-172(1997).
CC -1- FUNCTION: Synthesis of alpha-(+)-lipoic acid. It may be involved
CC in the sulfur insertion chemistry in lipoate biosynthesis (By
CC similarity).
CC -1- PATHWAY: Lipoate biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the biotin and lipoic acid synthetases
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y11708; CAA72400.1; -.
DR HAMAP; MF 00206; -.
DR InterPro; IPR006638; Elp3.
DR InterPro; IPR003698; Lipoate synth.
DR Pfam; PF04055; Radical_SAM.
DR SMART; SM00729; Elp3; I.
DR TIGRFAMs; TIGR00510; lipA; 1.
DR Iron-sulfur.
KW METAL 87 81 IRON-SULFUR (POTENTIAL).
FT METAL 91 91 IRON-SULFUR (POTENTIAL).
FT METAL 94 94 IRON-SULFUR (POTENTIAL).
SQ SEQUENCE 322 AA; 36297 MW; 45E7185215486847 CRC64;

Query Match 80.9%; Score 38; DB 1; Length 322;
Best Local Similarity 70.0%; Pred. No. 3.1;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10
DB 188 TVAGNYLTIR 197
```

```
RESULT 3
LIPA AGRTS STANDARD; PRT; 323 AA.
AC Q08UG1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lipoic acid synthetase (Lip-syn) (Lipoate synthase).
GN LIPA OR ATU1436 OR AGR_C_2846.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chena Y., Paulsen I.T., Bisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Doan M.,
RA Chumley F., Tingey S.V., Tomb J.-P., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
CC C58".
CC Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Ourillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Roumel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
CC Agrobacterium tumefaciens C58".
CC Science 294:2323-2328(2001).
CC -1- FUNCTION: Synthesis of alpha-(+)-lipoic acid. It may be involved
CC in the sulfur insertion chemistry in lipoate biosynthesis (By
CC similarity).
CC -1- PATHWAY: lipoate biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: Belongs to the biotin and lipoic acid synthetases
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE009104; AAL42442.1; -.
DR EMBL; AE008069; AAKG7227.1; -.
DR PIR; AD2753; AD2753.
DR PIR; B97534; B97534.
DR HAMAP; MF 00206; -.
DR InterPro; IPR006638; Elp3.
DR InterPro; IPR003698; Lipoate synth.
DR InterPro; IPR007197; Radical_SAM.
DR Pfam; PF04055; Radical_SAM; 1.
DR SMART; SM00729; Elp3; I.
DR TIGRFAMs; TIGR00510; lipA; 1.
DR Iron-sulfur; Complete proteome.
KW METAL 87 87 IRON-SULFUR (POTENTIAL).
FT METAL 91 91 IRON-SULFUR (POTENTIAL).
FT METAL 94 94 IRON-SULFUR (POTENTIAL).
SQ SEQUENCE 323 AA; 36340 MW; 74B294773BC784D2 CRC64;
```

Query Match 76.6%; Score 36; DB 1; Length 323;  
Best Local Similarity 70.0%; Pred. No. 8;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10  
Db 189 TVPGNYLTVR 198  
||| ||| |||

RESULT 4  
ID\_FABA\_PIG STANDARD; PRT; 131 AA.  
AC O97788;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Fatty acid-binding protein, adipocyte (AFABP) (Adipocyte lipid-binding protein) (ALBP) (A-FABP) (AP2).  
GN FABP4.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OC NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RC STRAIN=Duroc; TISSUE=Liver;  
RX MEDLINE=9909248; PubMed=980671;  
RA Gerbens F.N.A., Jansen A., van Erp A.J.M., Harders F., Meuwissen T.H.E., Rettenberger G.F.W., Veerkamp J.H., te Pas M.F.W.;  
RT "The adipocyte fatty acid-binding protein locus: characterization and association with intramuscular fat content in pigs."  
RT Mamm. Genome 9:1022-1026(1998).  
RN [2]  
RP SEQUENCE OF 36-47, AND TISSUE SPECIFICITY.  
RC TISSUE=Fat;  
RX MEDLINE=90241143; PubMed=2334399;  
RA Armstrong M.X., Bernlohr D.A., Storch J., Clarke S.D.;  
RT "The purification and characterization of a fatty acid binding protein specific to pig (Sus domesticus) adipose tissue."  
RL Biochem. J. 267:373-378(1990).  
CC -!- FUNCTION: Lipid transport protein in adipocytes. Binds both long chain fatty acid and retinoic acid. Involved in the regulation of intramuscular fat accretion.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
CC -!- TISSUE SPECIFICITY: Adipose tissue.  
CC -!- PM: The N-terminus is blocked.  
CC -!- SIMILARITY: Belongs to the fatty-acid binding protein (FABP) family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; Y16039; CAA75995.1; -;  
EMBL; AJ416020; CAC35166.1; -;  
DR HSPF; P04117; ILID.  
DR InterPro; IPR000463; Fatty acid BP.  
DR InterPro; IPR000566; Lipocln\_cytFABP.  
DR Pfam; PF00061; lipocalin; 1.  
DR PRINTS; PR00178; FATTYACIDBP.  
DR PROSITE; PS00214; FABP; 1.  
KW Transport; Lipid-binding; Phosphorylation.  
FT INIT\_MET 0  
FT MOD\_RES 19 0  
FT PHOSPHORYLATION (BY TYR-KINASES)  
FT (BY SIMILARITY).  
FT SEQUENCE 131 AA; 14545 MW; 656CB0CA08CD4B2 CRC64;

Query Match 72.3%; Score 34; DB 1; Length 131;

Best Local Similarity 60.0%; Pred. No. 8;  
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TVSGNLTIR 10  
Db 43 TVNGDMITIR 52  
||| ||| |||

RESULT 5  
ID\_RL6\_BACST STANDARD; PRT; 177 AA.  
AC P02391;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE 50S ribosomal protein L6 (BL10).  
GN RPLP.  
OS Bacillus stearothermophilus.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.  
OC NCBI\_TaxID=1422;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91093287; PubMed=1985969;  
RA Ramakrishnan V., Gerschman S.E.;  
RT "Cloning, sequencing, and overexpression of genes for ribosomal proteins from Bacillus stearothermophilus."  
RL J. Biol. Chem. 266:880-885(1991).  
RN [2]  
RP SEQUENCE.  
RA Kimura M., Rawlings N., Appelt K.;  
RT "The amino acid sequence of protein BL10 from the 50S subunit of the Bacillus stearothermophilus ribosome."  
RL FEBS Lett. 136:58-64(1981).  
RN [3]  
RP SEQUENCE OF 149-163, AND CROSS-LINKING TO RRNA.  
RC STRAIN=799;  
RX MEDLINE=96003638; PubMed=7556101;  
RA Urlaub H., Kruft V., Bischof O., Mueller E.-C., Wittmann-Liebold B.;  
RT "Protein-rRNA binding features and their structural and functional implications in ribosomes as determined by cross-linking studies."  
RL EMBO J. 14:4578-4588(1995).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).  
RX MEDLINE=94085364; PubMed=8262035;  
RA Golden B.L., Ramakrishnan V., White S.W.;  
RT "Ribosomal protein L6: structural evidence of gene duplication from a primitive RNA binding protein."  
RL EMBO J. 12:4901-4908(1993).  
CC -!- FUNCTION: This protein binds directly to 23S ribosomal RNA and is located at the aminoacyl-tRNA binding site of the peptidyltransferase center.  
CC -!- SIMILARITY: Belongs to the L6p family of ribosomal proteins.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; M57622; AAA22700.1; ALT\_SEQ.  
DR PIR; A02766; R3BSOP.  
DR PDB; 1RL6; 02-FEB-99.  
DR PDB; 1C04; 05-NOV-99.  
DR InterPro; IPR000702; Ribosomal L6.  
DR InterPro; IPR002358; Ribosomal\_L6\_1.  
DR Pfam; PF00347; Ribosomal\_L6; 2.  
DR PRINTS; PR00059; RIBOSOMALL6.  
DR PRODOM; PD002236; Ribosomal\_L6; 1.  
DR PROSITE; PS00525; RIBOSOMAL\_L6; 1.  
KW Ribosomal protein; rRNA-binding; 3D-structure.  
FT INIT\_MET 0

```

FT STRAND      8      9
FT TURN        12     13
FT STRAND      15     19
FT TURN        20     21
FT STRAND      22     27
FT TURN        28     29
FT STRAND      30     35
FT TURN        38     39
FT STRAND      41     45
FT TURN        46     47
FT STRAND      48     52
FT TURN        58     79
FT STRAND      80     80
FT TURN        82     89
FT STRAND      91     92
FT TURN        94     98
FT STRAND      99    100
FT STRAND     101    105
FT TURN     112    114
FT STRAND     118    119
FT TURN     120    126
FT STRAND     127    128
FT STRAND     129    134
FT TURN     137    148
FT STRAND     149    150
FT TURN     155    157
FT STRAND     160    163
FT TURN     164    165
SQ SEQUENCE    177 AA; 19168 MW; B4EC662A419616B3 CRC64;

Query Match      72.3%; Score 34; DB 1; Length 177;
Best Local Similarity 55.6%; Pred. No. 11;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 TVSGNLT 9
DB      43 TVEGNVITV 51

RESULT 6
IF5 SCHPO
ID ID5 SCHPO STANDARD; PRT; 244 AA.
AC Q94476;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Eukaryotic translation initiation factor 6 (eIF-6).
GN TIP6 OR SPCC1919.09.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moutle S., Mungall K., Murphy L., Niblett D., Odeh C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkert G., Aert R., Robben J., Grymonprez B.,
RA Waltens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer S., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,

```

---

```

RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RL "The genome sequence of Schizosaccharomyces pombe.";
CC Nature 415:871-880(2002).
CC -!- FUNCTION: Binds to the 60S ribosomal subunit and prevents its
CC association with the 40S ribosomal subunit to form the 80S
CC initiation complex (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the eIF-6 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL035075; CAA22640.1; -
CC PIR; T41234; T41234.
CC HSP; Q12522; 1G62.
CC GeneDB_SPombe; SPCC1919.09; -.
CC InterPro; IPR002769; eIF6.
CC Pfam; PF01912; eIF6; 1.
CC ProDom; PD006880; eIF6; 1.
CC SMART; SM00654; eIF6; 1.
KW Initiation factor; Protein biosynthesis; Phosphorylation.
FT MOD_RES 174 174
FT MOD_RES 175 175
FT MOD_RES 176 176
FT MOD_RES 177 177
FT MOD_RES 178 178
FT MOD_RES 179 179
FT MOD_RES 180 180
FT MOD_RES 181 181
FT MOD_RES 182 182
FT MOD_RES 183 183
FT MOD_RES 184 184
FT MOD_RES 185 185
FT MOD_RES 186 186
FT MOD_RES 187 187
FT MOD_RES 188 188
FT MOD_RES 189 189
FT MOD_RES 190 190
FT MOD_RES 191 191
FT MOD_RES 192 192
FT MOD_RES 193 193
FT MOD_RES 194 194
FT MOD_RES 195 195
FT MOD_RES 196 196
FT MOD_RES 197 197
FT MOD_RES 198 198
FT MOD_RES 199 199
FT MOD_RES 200 200
FT MOD_RES 201 201
FT MOD_RES 202 202
FT MOD_RES 203 203
FT MOD_RES 204 204
FT MOD_RES 205 205
FT MOD_RES 206 206
FT MOD_RES 207 207
FT MOD_RES 208 208
FT MOD_RES 209 209
FT MOD_RES 210 210
FT MOD_RES 211 211
FT MOD_RES 212 212
FT MOD_RES 213 213
FT MOD_RES 214 214
FT MOD_RES 215 215
FT MOD_RES 216 216
FT MOD_RES 217 217
FT MOD_RES 218 218
FT MOD_RES 219 219
FT MOD_RES 220 220
FT MOD_RES 221 221
FT MOD_RES 222 222
FT MOD_RES 223 223
FT MOD_RES 224 224
FT MOD_RES 225 225
FT MOD_RES 226 226
FT MOD_RES 227 227
FT MOD_RES 228 228
FT MOD_RES 229 229
FT MOD_RES 230 230
FT MOD_RES 231 231
FT MOD_RES 232 232
FT MOD_RES 233 233
FT MOD_RES 234 234
FT MOD_RES 235 235
FT MOD_RES 236 236
FT MOD_RES 237 237
FT MOD_RES 238 238
FT MOD_RES 239 239
FT MOD_RES 240 240
FT MOD_RES 241 241
FT MOD_RES 242 242
FT MOD_RES 243 243
FT MOD_RES 244 244
FT MOD_RES 245 245
FT MOD_RES 246 246
FT MOD_RES 247 247
FT MOD_RES 248 248
FT MOD_RES 249 249
FT MOD_RES 250 250
FT MOD_RES 251 251
FT MOD_RES 252 252
FT MOD_RES 253 253
FT MOD_RES 254 254
FT MOD_RES 255 255
FT MOD_RES 256 256
FT MOD_RES 257 257
FT MOD_RES 258 258
FT MOD_RES 259 259
FT MOD_RES 260 260
FT MOD_RES 261 261
FT MOD_RES 262 262
FT MOD_RES 263 263
FT MOD_RES 264 264
FT MOD_RES 265 265
FT MOD_RES 266 266
FT MOD_RES 267 267
FT MOD_RES 268 268
FT MOD_RES 269 269
FT MOD_RES 270 270
FT MOD_RES 271 271
FT MOD_RES 272 272
FT MOD_RES 273 273
FT MOD_RES 274 274
FT MOD_RES 275 275
FT MOD_RES 276 276
FT MOD_RES 277 277
FT MOD_RES 278 278
FT MOD_RES 279 279
FT MOD_RES 280 280
FT MOD_RES 281 281
FT MOD_RES 282 282
FT MOD_RES 283 283
FT MOD_RES 284 284
FT MOD_RES 285 285
FT MOD_RES 286 286
FT MOD_RES 287 287
FT MOD_RES 288 288
FT MOD_RES 289 289
FT MOD_RES 290 290
FT MOD_RES 291 291
FT MOD_RES 292 292
FT MOD_RES 293 293
FT MOD_RES 294 294
FT MOD_RES 295 295
FT MOD_RES 296 296
FT MOD_RES 297 297
FT MOD_RES 298 298
FT MOD_RES 299 299
FT MOD_RES 300 300
FT MOD_RES 301 301
FT MOD_RES 302 302
FT MOD_RES 303 303
FT MOD_RES 304 304
FT MOD_RES 305 305
FT MOD_RES 306 306
FT MOD_RES 307 307
FT MOD_RES 308 308
FT MOD_RES 309 309
FT MOD_RES 310 310
FT MOD_RES 311 311
FT MOD_RES 312 312
FT MOD_RES 313 313
FT MOD_RES 314 314
FT MOD_RES 315 315
FT MOD_RES 316 316
FT MOD_RES 317 317
FT MOD_RES 318 318
FT MOD_RES 319 319
FT MOD_RES 320 320
FT MOD_RES 321 321
FT MOD_RES 322 322
FT MOD_RES 323 323
FT MOD_RES 324 324
FT MOD_RES 325 325
FT MOD_RES 326 326
FT MOD_RES 327 327
FT MOD_RES 328 328
FT MOD_RES 329 329
FT MOD_RES 330 330
FT MOD_RES 331 331
FT MOD_RES 332 332
FT MOD_RES 333 333
FT MOD_RES 334 334
FT MOD_RES 335 335
FT MOD_RES 336 336
FT MOD_RES 337 337
FT MOD_RES 338 338
FT MOD_RES 339 339
FT MOD_RES 340 340
FT MOD_RES 341 341
FT MOD_RES 342 342
FT MOD_RES 343 343
FT MOD_RES 344 344
FT MOD_RES 345 345
FT MOD_RES 346 346
FT MOD_RES 347 347
FT MOD_RES 348 348
FT MOD_RES 349 349
FT MOD_RES 350 350
FT MOD_RES 351 351
FT MOD_RES 352 352
FT MOD_RES 353 353
FT MOD_RES 354 354
FT MOD_RES 355 355
FT MOD_RES 356 356
FT MOD_RES 357 357
FT MOD_RES 358 358
FT MOD_RES 359 359
FT MOD_RES 360 360
FT MOD_RES 361 361
FT MOD_RES 362 362
FT MOD_RES 363 363
FT MOD_RES 364 364
FT MOD_RES 365 365
FT MOD_RES 366 366
FT MOD_RES 367 367
FT MOD_RES 368 368
FT MOD_RES 369 369
FT MOD_RES 370 370
FT MOD_RES 371 371
FT MOD_RES 372 372
FT MOD_RES 373 373
FT MOD_RES 374 374
FT MOD_RES 375 375
FT MOD_RES 376 376
FT MOD_RES 377 377
FT MOD_RES 378 378
FT MOD_RES 379 379
FT MOD_RES 380 380
FT MOD_RES 381 381
FT MOD_RES 382 382
FT MOD_RES 383 383
FT MOD_RES 384 384
FT MOD_RES 385 385
FT MOD_RES 386 386
FT MOD_RES 387 387
FT MOD_RES 388 388
FT MOD_RES 389 389
FT MOD_RES 390 390
FT MOD_RES 391 391
FT MOD_RES 392 392
FT MOD_RES 393 393
FT MOD_RES 394 394
FT MOD_RES 395 395
FT MOD_RES 396 396
FT MOD_RES 397 397
FT MOD_RES 398 398
FT MOD_RES 399 399
FT MOD_RES 400 400
FT MOD_RES 401 401
FT MOD_RES 402 402
FT MOD_RES 403 403
FT MOD_RES 404 404
FT MOD_RES 405 405
FT MOD_RES 406 406
FT MOD_RES 407 407
FT MOD_RES 408 408
FT MOD_RES 409 409
FT MOD_RES 410 410
FT MOD_RES 411 411
FT MOD_RES 412 412
FT MOD_RES 413 413
FT MOD_RES 414 414
FT MOD_RES 415 415
FT MOD_RES 416 416
FT MOD_RES 417 417
FT MOD_RES 418 418
FT MOD_RES 419 419
FT MOD_RES 420 420
FT MOD_RES 421 421
FT MOD_RES 422 422
FT MOD_RES 423 423
FT MOD_RES 424 424
FT MOD_RES 425 425
FT MOD_RES 426 426
FT MOD_RES 427 427
FT MOD_RES 428 428
FT MOD_RES 429 429
FT MOD_RES 430 430
FT MOD_RES 431 431
FT MOD_RES 432 432
FT MOD_RES 433 433
FT MOD_RES 434 434
FT MOD_RES 435 435
FT MOD_RES 436 436
FT MOD_RES 437 437
FT MOD_RES 438 438
FT MOD_RES 439 439
FT MOD_RES 440 440
FT MOD_RES 441 441
FT MOD_RES 442 442
FT MOD_RES 443 443
FT MOD_RES 444 444
FT MOD_RES 445 445
FT MOD_RES 446 446
FT MOD_RES 447 447
FT MOD_RES 448 448
FT MOD_RES 449 449
FT MOD_RES 450 450
FT MOD_RES 451 451
FT MOD_RES 452 452
FT MOD_RES 453 453
FT MOD_RES 454 454
FT MOD_RES 455 455
FT MOD_RES 456 456
FT MOD_RES 457 457
FT MOD_RES 458 458
FT MOD_RES 459 459
FT MOD_RES 460 460
FT MOD_RES 461 461
FT MOD_RES 462 462
FT MOD_RES 463 463
FT MOD_RES 464 464
FT MOD_RES 465 465
FT MOD_RES 466 466
FT MOD_RES 467 467
FT MOD_RES 468 468
FT MOD_RES 469 469
FT MOD_RES 470 470
FT MOD_RES 471 471
FT MOD_RES 472 472
FT MOD_RES 473 473
FT MOD_RES 474 474
FT MOD_RES 475 475
FT MOD_RES 476 476
FT MOD_RES 477 477
FT MOD_RES 478 478
FT MOD_RES 479 479
FT MOD_RES 480 480
FT MOD_RES 481 481
FT MOD_RES 482 482
FT MOD_RES 483 483
FT MOD_RES 484 484
FT MOD_RES 485 485
FT MOD_RES 486 486
FT MOD_RES 487 487
FT MOD_RES 488 488
FT MOD_RES 489 489
FT MOD_RES 490 490
FT MOD_RES 491 491
FT MOD_RES 492 492
FT MOD_RES 493 493
FT MOD_RES 494 494
FT MOD_RES 495 495
FT MOD_RES 496 496
FT MOD_RES 497 497
FT MOD_RES 498 498
FT MOD_RES 499 499
FT MOD_RES 500 500
FT MOD_RES 501 501
FT MOD_RES 502 502
FT MOD_RES 503 503
FT MOD_RES 504 504
FT MOD_RES 505 505
FT MOD_RES 506 506
FT MOD_RES 507 507
FT MOD_RES 508 508
FT MOD_RES 509 509
FT MOD_RES 510 510
FT MOD_RES 511 511
FT MOD_RES 512 512
FT MOD_RES 513 513
FT MOD_RES 514 514
FT MOD_RES 515 515
FT MOD_RES 516 516
FT MOD_RES 517 517
FT MOD_RES 518 518
FT MOD_RES 519 519
FT MOD_RES 520 520
FT MOD_RES 521 521
FT MOD_RES 522 522
FT MOD_RES 523 523
FT MOD_RES 524 524
FT MOD_RES 525 525
FT MOD_RES 526 526
FT MOD_RES 527 527
FT MOD_RES 528 528
FT MOD_RES 529 529
FT MOD_RES 530 530
FT MOD_RES 531 531
FT MOD_RES 532 532
FT MOD_RES 533 533
FT MOD_RES 534 534
FT MOD_RES 535 535
FT MOD_RES 536 536
FT MOD_RES 537 537
FT MOD_RES 538 538
FT MOD_RES 539 539
FT MOD_RES 540 540
FT MOD_RES 541 541
FT MOD_RES 542 542
FT MOD_RES 543 543
FT MOD_RES 544 544
FT MOD_RES 545 545
FT MOD_RES 546 546
FT MOD_RES 547 547
FT MOD_RES 548 548
FT MOD_RES 549 549
FT MOD_RES 550 550
FT MOD_RES 551 551
FT MOD_RES 552 552
FT MOD_RES 553 553
FT MOD_RES 554 554
FT MOD_RES 555 555
FT MOD_RES 556 556
FT MOD_RES 557 557
FT MOD_RES 558 558
FT MOD_RES 559 559
FT MOD_RES 560 560
FT MOD_RES 561 561
FT MOD_RES 562 562
FT MOD_RES 563 563
FT MOD_RES 564 564
FT MOD_RES 565 565
FT MOD_RES 566 566
FT MOD_RES 567 567
FT MOD_RES 568 568
FT MOD_RES 569 569
FT MOD_RES 570 570
FT MOD_RES 571 571
FT MOD_RES 572 572
FT MOD_RES 573 573
FT MOD_RES 574 574
FT MOD_RES 575 575
FT MOD_RES 576 576
FT MOD_RES 577 577
FT MOD_RES 578 578
FT MOD_RES 579 579
FT MOD_RES 580 580
FT MOD_RES 581 581
FT MOD_RES 582 582
FT MOD_RES 583 583
FT MOD_RES 584 584
FT MOD_RES 585 585
FT MOD_RES 586 586
FT MOD_RES 587 587
FT MOD_RES 588 588
FT MOD_RES 589 589
FT MOD_RES 590 590
FT MOD_RES 591 591
FT MOD_RES 592 592
FT MOD_RES 593 593
FT MOD_RES 594 594
FT MOD_RES 595 595
FT MOD_RES 596 596
FT MOD_RES 597 597
FT MOD_RES 598 598
FT MOD_RES 599 599
FT MOD_RES 600 600
FT MOD_RES 601 601
FT MOD_RES 602 602
FT MOD_RES 603 603
FT MOD_RES 604 604
FT MOD_RES 605 605
FT MOD_RES 606 606
FT MOD_RES 607 607
FT MOD_RES 608 608
FT MOD_RES 609 609
FT MOD_RES 610 610
FT MOD_RES 611 611
FT MOD_RES 612 612
FT MOD_RES 613 613
FT MOD_RES 614 614
FT MOD_RES 615 615
FT MOD_RES 616 616
FT MOD_RES 617 617
FT MOD_RES 618 618
FT MOD_RES 619 619
FT MOD_RES 620 620
FT MOD_RES 621 621
FT MOD_RES 622 622
FT MOD_RES 623 623
FT MOD_RES 624 624
FT MOD_RES 625 625
FT MOD_RES 626 626
FT MOD_RES 627 627
FT MOD_RES 628 628
FT MOD_RES 629 629
FT MOD_RES 630 630
FT MOD_RES 631 631
FT MOD_RES 632 632
FT MOD_RES 633 633
FT MOD_RES 634 634
FT MOD_RES 635 635
FT MOD_RES 636 636
FT MOD_RES 637 637
FT MOD_RES 638 638
FT MOD_RES 639 639
FT MOD_RES 640 640
FT MOD_RES 641 641
FT MOD_RES 642 642
FT MOD_RES 643 643
FT MOD_RES 644 644
FT MOD_RES 645 645
FT MOD_RES 646 646
FT MOD_RES 647 647
FT MOD_RES 648 648
FT MOD_RES 649 649
FT MOD_RES 650 650
FT MOD_RES 651 651
FT MOD_RES 652 652
FT MOD_RES 653 653
FT MOD_RES 654 654
FT MOD_RES 655 655
FT MOD_RES 656 656
FT MOD_RES 657 657
FT MOD_RES 658 658
FT MOD_RES 659 659
FT MOD_RES 660 660
FT MOD_RES 661 661
FT MOD_RES 662 662
FT MOD_RES 663 663
FT MOD_RES 664 664
FT MOD_RES 665 665
FT MOD_RES 666 666
FT MOD_RES 667 667
FT MOD_RES 668 668
FT MOD_RES 669 669
FT MOD_RES 670 670
FT MOD_RES 671 671
FT MOD_RES 672 672
FT MOD_RES 673 673
FT MOD_RES 674 674
FT MOD_RES 675 675
FT MOD_RES 676 676
FT MOD_RES 677 677
FT MOD_RES 678 678
FT MOD_RES 679 679
FT MOD_RES 680 680
FT MOD_RES 681 681
FT MOD_RES 682 682
FT MOD_RES 683 683
FT MOD_RES 684 684
FT MOD_RES 685 685
FT MOD_RES 686 686
FT MOD_RES 687 687
FT MOD_RES 688 688
FT MOD_RES 689 689
FT MOD_RES 690 690
FT MOD_RES 691 691
FT MOD_RES 692 692
FT MOD_RES 693 693
FT MOD_RES 694 694
FT MOD_RES 695 695
FT MOD_RES 696 696
FT MOD_RES 697 697
FT MOD_RES 698 698
FT MOD_RES 699 699
FT MOD_RES 700 700
FT MOD_RES 701 701
FT MOD_RES 702 702
FT MOD_RES 703 703
FT MOD_RES 704 704
FT MOD_RES 705 705
FT MOD_RES 706 706
FT MOD_RES 707 707
FT MOD_RES 708 708
FT MOD_RES 709 709
FT MOD_RES 710 710
FT MOD_RES 711 711
FT MOD_RES 712 712
FT MOD_RES 713 713
FT MOD_RES 714 714
FT MOD_RES 715 715
FT MOD_RES 716 716
FT MOD_RES 717 717
FT MOD_RES 718 718
FT MOD_RES 719 719
FT MOD_RES 720 720
FT MOD_RES 721 721
FT MOD_RES 722 722
FT MOD_RES 723 723
FT MOD_RES 724 724
FT MOD_RES 725 725
FT MOD_RES 726 726
FT MOD_RES 727 727
FT MOD_RES 728 728
FT MOD_RES 729 729
FT MOD_RES 730 730
FT MOD_RES 731 731
FT MOD_RES 732 732
FT MOD_RES 733 733
FT MOD_RES 734 734
FT MOD_RES 735 735
FT MOD_RES 736 736
FT MOD_RES 737 737
FT MOD_RES 738 738
FT MOD_RES 739 739
FT MOD_RES 740 740
FT MOD_RES 741 741
FT MOD_RES 742 742
FT MOD_RES 743 743
FT MOD_RES 744 744
FT MOD_RES 745 745
FT MOD_RES 746 746
FT MOD_RES 747 747
FT MOD_RES 748 748
FT MOD_RES 749 749
FT MOD_RES 750 750
FT MOD_RES 751 751
FT MOD_RES 752 752
FT MOD_RES 753 753
FT MOD_RES 754 754
FT MOD_RES 755 755
FT MOD_RES 756 756
FT MOD_RES 757 757
FT MOD_RES 758 758
FT MOD_RES 759 759
FT MOD_RES 760 760
FT MOD_RES 761 761
FT MOD_RES 762 762
FT MOD_RES 763 763
FT MOD_RES 764 764
FT MOD_RES 765 765
FT MOD_RES 766 766
FT MOD_RES 767 767
FT MOD_RES 768 768
FT MOD_RES 769 769
FT MOD_RES 770 770
FT MOD_RES 771 771
FT MOD_RES 772 772
FT MOD_RES 773 773
FT MOD_RES 774 774
FT MOD_RES 775 775
FT MOD_RES 776 776
FT MOD_RES 777 777
FT MOD_RES 778 778
FT MOD_RES 779 779
FT MOD_RES 780 780
FT MOD_RES 781 781
FT MOD_RES 782 782
FT MOD_RES 783 783
FT MOD_RES 784 784
FT MOD_RES 785 785
FT MOD_RES 786 786
FT MOD_RES 787 787
FT MOD_RES 788 788
FT MOD_RES 789 789
FT MOD_RES 790 790
FT MOD_RES 791 791
FT MOD_RES 792 792
FT MOD_RES 793 793
FT MOD_RES 794 794
FT MOD_RES 795 795
FT MOD_RES 796 796
FT MOD_RES 797 797
FT MOD_RES 798 798
FT MOD_RES 799 799
FT MOD_RES 800 800
FT MOD_RES 801 801
FT MOD_RES 802 802
FT MOD_RES 803 803
FT MOD_RES 804 804
FT MOD_RES 805 805
FT MOD_RES 806 806
FT MOD_RES 807 807
FT MOD_RES 808 808
FT MOD_RES 809 809
FT MOD_RES 810 810
FT MOD_RES 811 811
FT MOD_RES 812 812
FT MOD_RES 813 813
FT MOD_RES 814 814
FT MOD_RES 815 815
FT MOD_RES 816 816
FT MOD_RES 817 817
FT MOD_RES 818 818
FT MOD_RES 819 819
FT MOD_RES 820 820
FT MOD_RES 821 821
FT MOD_RES 822 822
FT MOD_RES 823 823
FT MOD_RES 824 824
FT MOD_RES 825 825
FT MOD_RES 826 826
FT MOD_RES 827 827
FT MOD_RES 828 828
FT MOD_RES 829 829
FT MOD_RES 830 830
FT MOD_RES 831 831
FT MOD_RES 832 832
FT MOD_RES 833 833
FT MOD_RES 834 834
FT MOD_RES 835 835
FT MOD_RES 836 836
FT MOD_RES 837 837
FT MOD_RES 838 838
FT MOD_RES 839 839
FT MOD_RES 840 840
FT MOD_RES 841 841
FT MOD_RES 842 842
FT MOD_RES 843 843
FT MOD_RES 844 844
FT MOD_RES 845 845
FT MOD_RES 846 846
FT MOD_RES 847 847
FT MOD_RES 848 848
FT MOD_RES 849 849
FT MOD_RES 850 850
FT MOD_RES 851 851
FT MOD_RES 852 852
FT MOD_RES 853 853
FT MOD_RES 854 854
FT MOD_RES 855 855
FT MOD_RES 856 856
FT MOD_RES 857 857
FT MOD_RES 858 858
FT MOD_RES 859 859
FT MOD_RES 860 860
FT MOD_RES 861 861
FT MOD_RES 862 862
FT MOD_RES 863 863
FT MOD_RES 864 864
FT MOD_RES 865 865
FT MOD_RES 866 866
FT MOD_RES 867 867
FT MOD_RES 868 868
FT MOD_RES 869 869
FT MOD_RES 870 870
FT MOD_RES 871 871
FT MOD_RES 872 872
FT MOD_RES 873 873
FT MOD_RES 874 874
FT MOD_RES 875 875
FT MOD_RES 876 876
FT MOD_RES 877 877
FT MOD_RES 878 878
FT MOD_RES 879 879
FT MOD_RES 880 880
FT MOD_RES 881 881
FT MOD_RES 882 882
FT MOD_RES 883 883
FT MOD_RES 884 884
FT MOD_RES 885 885
FT MOD_RES 886 886
FT MOD_RES 887 887
FT MOD_RES 888 888
FT MOD_RES 889 889
FT MOD_RES 890 890
FT MOD_RES 891 891
FT MOD_RES 892 892
FT MOD_RES 893 893
FT MOD_RES 894 894
FT MOD_RES 895 895
FT MOD_RES 896 896
FT MOD_RES 897 897
FT MOD_RES 898 898
FT MOD_RES 899 899
FT MOD_RES 900 900
FT MOD_RES 901 901
FT MOD_RES 902 902
FT MOD_RES 903 903
FT MOD_RES 904 904
FT MOD_RES 905 905
FT MOD_RES 906 906
FT MOD_RES 907 907
FT MOD_RES 908 908
FT MOD_RES 909 909
FT MOD_RES 910 910
FT MOD_RES 911 911
FT MOD_RES 912 912
FT MOD_RES 913 913
FT MOD_RES 914 914
FT MOD_RES 915 915
FT MOD_RES 916 916
FT MOD_RES 917 917
FT MOD_RES 918 918
FT MOD_RES 919 919
FT MOD_RES 920 920
FT MOD_RES 921 921
FT MOD_RES 922 922
FT MOD_RES 923 923
FT MOD_RES 924 924
FT MOD_RES 925 925
FT MOD_RES 926 926
FT MOD_RES 927 927
FT MOD_RES 928 928
FT MOD_RES 929 929
FT MOD_RES 930 930
FT MOD_RES 931 931
FT MOD_RES 932 932
FT MOD_RES 933 933
FT MOD_RES 934 934
FT MOD_RES 935 935
FT MOD_RES 936 936
FT MOD_RES 937 937
FT MOD_RES 938 938
FT MOD_RES 939 939
FT MOD_RES 940 940
FT MOD_RES 941 941
FT MOD_RES 942 942
FT MOD_RES 943 943
FT MOD_RES 944 944
FT MOD_RES 945 945
FT MOD_RES 946 946
FT MOD_RES 947 947
FT MOD_RES 948 948
FT MOD_RES 949 949
FT MOD_RES 950 950
FT MOD_RES 951 951
FT MOD_RES 952 952
FT MOD_RES 953 953
FT MOD_RES 954 954
FT MOD_RES 955 955
FT MOD_RES 956 956
FT MOD_RES 957 957
FT MOD_RES 958 958
FT MOD_RES 959 959
FT MOD_RES 960 960
FT MOD_RES 961 961
FT MOD_RES 962 962
FT MOD_RES 963 963
FT MOD_RES 964 964
FT MOD_RES 965 965
FT MOD_RES 966 966
FT MOD_RES 967 967
FT MOD_RES 968 968
FT MOD_RES 969 969
FT MOD_RES 970 970
FT MOD_RES 971 971
FT MOD_RES 972 972
FT MOD_RES 973 973
FT MOD_RES 974 974
FT MOD_RES 975 975
FT MOD_RES 976 976
FT MOD_RES 977 977
FT MOD_RES 978 978
FT MOD_RES 979 979
FT MOD_RES 980 980
FT MOD_RES 981 981
FT MOD_RES 982 982
FT MOD_RES 983 983
FT MOD_RES 984 984
FT MOD_RES 985 985
FT MOD_RES 986 986
FT MOD_RES 987 987
FT MOD_RES 988 988
FT MOD_RES 989 989
FT MOD_RES 990 990
FT MOD_RES 991 991
FT MOD_RES 992 992
FT MOD_RES 993 993
FT MOD_RES 994 994
FT MOD_RES 995 995
FT MOD_RES 996 996
FT MOD_RES 997 997
FT MOD_RES 998 998
FT MOD_RES 999 999
FT MOD_RES 1000 1000
FT MOD_RES 1001 1001
FT MOD_RES 1002 1002
FT MOD_RES 1003 1003
FT MOD_RES 1004 1004
FT MOD_RES 1005 1005
FT MOD_RES 1006 1006
FT MOD_RES 1007 1007
FT MOD_RES 1008 1008
FT MOD_RES 1009 1009
FT MOD_RES 1010 1010
FT MOD_RES 1011 1011
FT MOD_RES 1012 1012
FT MOD_RES 1013 1013
FT MOD_RES 1014 1014
FT MOD_RES 1015 1015
FT MOD_RES 1016 1016
FT MOD_RES 1017 1017
FT MOD_RES 1018 1018
FT MOD_RES 1019 1019
FT MOD_RES 1020 1020
FT MOD_RES 1021 1021
FT MOD_RES 1022 1022
FT MOD_RES 1023 1023
FT MOD_RES 1024 1024
FT MOD_RES 1025 1025
FT MOD_RES 1026 1026
FT MOD_RES 1027 1027
FT MOD_RES 1028 1028
FT MOD_RES 1029 1029
FT MOD_RES 1030 1030
FT MOD_RES 1031 1031
FT MOD_RES 1032 1032
FT MOD_RES 1033 1033
FT MOD_RES 1034 1034
FT MOD_RES 1035 1035
FT MOD_RES 1036 1036
FT MOD_RES 1037 1037
FT MOD_RES 1038 1
```

RT olfactory receptor gene families";  
 RL Hum. Mol. Genet. 11:535-546(2002).  
 RN [3]  
 RP ERATUM.  
 RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tonnes-Priddy L.,  
 RA Trask B.J.;  
 RL Hum. Mol. Genet. 11:1683-1683(2002).  
 CC -!- FUNCTION: Potential odorant receptor.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AY073356; AAL61019.1; -;  
 DR EMBL; AY317610; AAP71001.1; -;  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm.1; 1.  
 DR PRINTS; PR00237; GPCRHOPOPSN.  
 DR PROSITE; PS00237; G-PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS0262; G-PROTEIN\_RECEP\_F1\_2; 1.  
 DR G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Multigene family; Olfaction.  
 FT DOMAIN 1 28 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 29 49 1 (POTENTIAL).  
 FT DOMAIN 50 57 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 58 78 2 (POTENTIAL).  
 FT DOMAIN 79 102 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 103 123 3 (POTENTIAL).  
 FT DOMAIN 124 136 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 137 157 4 (POTENTIAL).  
 FT DOMAIN 158 199 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 200 220 5 (POTENTIAL).  
 FT DOMAIN 221 240 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 241 261 6 (POTENTIAL).  
 FT DOMAIN 262 274 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 275 295 7 (POTENTIAL).  
 FT DOMAIN 296 314 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 100 192 BY SIMILARITY.  
 FT CARBOHYD 8 8 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 314 AA; 34549 MW; E143B2A1E0321BD9 CRC64;  
 Query Match 72.3%; Score 34; DB 1; Length 314;  
 Best Local Similarity 77.8%; Pred. No. 20;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 TVSGNLT 9  
 DB 41 TVSGNLTSTI 49  
 RESULT 8  
 OR25\_MOUSE STANDARD; PRT; 314 AA.  
 AC Q8VG07;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Putative olfactory receptor 204-10.  
 GN MOR204-10.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21676863; PubMed=11802173;  
 RZ Zhang X., Firestein S.,

RT "The olfactory receptor gene superfamily of the mouse";  
 RL Nat. Neurosci. 5:124-133(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21864068; PubMed=11875048;  
 RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tonnes-Priddy L.,  
 RA Trask B.J.;  
 RL Hum. Mol. Genet. 11:535-546(2002).  
 CC -!- FUNCTION: Potential odorant receptor.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AY073358; AAL61021.1; -;  
 DR EMBL; AY317602; AAP70997.1; -;  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm.1; 1.  
 DR PRINTS; PR00237; GPCRHOPOPSN.  
 DR PROSITE; PS00237; G-PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS0262; G-PROTEIN\_RECEP\_F1\_2; 1.  
 DR G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Multigene family; Olfaction.  
 FT DOMAIN 1 28 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 29 49 1 (POTENTIAL).  
 FT DOMAIN 50 57 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 58 78 2 (POTENTIAL).  
 FT DOMAIN 79 102 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 103 123 3 (POTENTIAL).  
 FT DOMAIN 124 136 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 137 157 4 (POTENTIAL).  
 FT DOMAIN 158 199 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 200 220 5 (POTENTIAL).  
 FT DOMAIN 221 240 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 241 261 6 (POTENTIAL).  
 FT DOMAIN 262 274 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 275 295 7 (POTENTIAL).  
 FT DOMAIN 296 314 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 100 192 BY SIMILARITY.  
 FT CARBOHYD 8 8 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 314 AA; 34658 MW; 71BC08397C5884BD CRC64;  
 Query Match 72.3%; Score 34; DB 1; Length 314;  
 Best Local Similarity 77.8%; Pred. No. 20;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 TVSGNLT 9  
 DB 41 TVSGNLTSTI 49  
 RESULT 9  
 OR28\_MOUSE STANDARD; PRT; 314 AA.  
 AC Q8VG04;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Putative olfactory receptor 204-13.  
 GN MOR204-13.

OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21676863; PubMed=11802173;  
 RA Zhang X., Firestein S.;  
 RA "The olfactory receptor gene superfamily of the mouse.";  
 RT Nat. Neurosci. 5:124-133(2002).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21864068; PubMed=11875048;  
 RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tonnes-Priddy L.,  
 RA Trask B.J.;  
 RA "Different evolutionary processes shaped the mouse and human  
 RT olfactory receptor gene families.";  
 RL Hum. Mol. Genet. 11:535-546(2002).  
 RN [3]  
 RN ERRATUM.  
 RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tonnes-Priddy L.,  
 RA Trask B.J.;  
 RA Hum. Mol. Genet. 11:1693-1693(2002).  
 CC -!- FUNCTION: Potential odorant receptor.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AY073361; AAL61024.1; -;  
 DR EMBL; AY317586; AAP70982.1; -;  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm.1; 1.  
 DR PRINTS; PR00237; GPCRHOPOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECP\_F1\_1; 1.  
 DR PROSITE; PS0262; G\_PROTEIN\_RECP\_F1\_2; 1.  
 DR G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Multigene family; Olfaction.  
 FT DOMAIN 1 28 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 29 49 1 (POTENTIAL).  
 FT DOMAIN 50 57 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 58 78 2 (POTENTIAL).  
 FT DOMAIN 79 102 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 103 123 3 (POTENTIAL).  
 FT DOMAIN 124 136 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 137 157 4 (POTENTIAL).  
 FT DOMAIN 158 199 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 200 220 5 (POTENTIAL).  
 FT DOMAIN 221 240 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 241 261 6 (POTENTIAL).  
 FT DOMAIN 262 274 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 275 295 7 (POTENTIAL).  
 FT DOMAIN 296 314 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 100 192 BY SIMILARITY.  
 FT CARBOHYD 8 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 314 AA; 34940 MW; 976122610A4A6EF CRC64;  
 Query Match 72.3%; Score 34; DB 1; Length 314;  
 Best Local Similarity 77.8%; Pred. No. 20;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qv 1 TVSGNLT 9  
 |||||:  
 Db 41 TVSGNLT 49

RESULT 10

OR32\_MOUSE  
 ID OR32\_MOUSE STANDARD; PRT; 314 AA.  
 AC Q8VFD2;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Putative olfactory receptor 204-17.  
 GN MOR204-17.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21676863; PubMed=11802173;  
 RA Zhang X., Firestein S.;  
 RA "The olfactory receptor gene superfamily of the mouse.";  
 RT Nat. Neurosci. 5:124-133(2002).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21864068; PubMed=11875048;  
 RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tonnes-Priddy L.,  
 RA Trask B.J.;  
 RA "Different evolutionary processes shaped the mouse and human  
 RT olfactory receptor gene families.";  
 RL Hum. Mol. Genet. 11:535-546(2002).  
 RN [3]  
 RN ERRATUM.  
 RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tonnes-Priddy L.,  
 RA Trask B.J.;  
 RA Hum. Mol. Genet. 11:1693-1693(2002).  
 CC -!- FUNCTION: Potential odorant receptor.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AY073599; AAL61262.1; -;  
 DR EMBL; AY317598; AAP70993.1; -;  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm.1; 1.  
 DR PRINTS; PR00237; GPCRHOPOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECP\_F1\_1; 1.  
 DR PROSITE; PS0262; G\_PROTEIN\_RECP\_F1\_2; 1.  
 DR G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Multigene family; Olfaction.  
 FT DOMAIN 1 28 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 29 49 1 (POTENTIAL).  
 FT DOMAIN 50 57 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 58 78 2 (POTENTIAL).  
 FT DOMAIN 79 102 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 103 123 3 (POTENTIAL).  
 FT DOMAIN 124 136 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 137 157 4 (POTENTIAL).  
 FT DOMAIN 158 199 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 200 220 5 (POTENTIAL).  
 FT DOMAIN 221 240 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 241 261 6 (POTENTIAL).  
 FT DOMAIN 262 274 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 275 295 7 (POTENTIAL).  
 FT DOMAIN 296 314 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 100 192 BY SIMILARITY.  
 FT CARBOHYD 8 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 314 AA; 34744 MW; 35FF7679DB37FD7E CRC64;  
 Query Match 72.3%; Score 34; DB 1; Length 314;  
 Best Local Similarity 77.8%; Pred. No. 20;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TVSGNULTI 9  
|||||: ||  
Db 41 TVSGNLSTI 49

RESULT 11  
OR33 MOUSE STANDARD; PRT; 314 AA.  
AC Q8VFD1;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Putative olfactory receptor 204-18.  
GN MOR204-18.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21676863; PubMed=11802173;  
RA Zhang X., Firestein S.;  
RT "The olfactory receptor gene superfamily of the mouse.";  
RL Nat. Neurosci. 5:124-133(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21864068; PubMed=11875048;  
RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tonnes-Priddy L.,  
RA Trask B.J.;  
RT "Different evolutionary processes shaped the mouse and human  
olfactory receptor gene families.";  
RL Hum. Mol. Genet. 11:535-546(2002).  
RN [3]  
RP ERRATUM.  
RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tonnes-Priddy L.,  
RA Trask B.J.;  
RL Hum. Mol. Genet. 11:1683-1683(2002).  
CC -!- FUNCTION: Potential odorant receptor.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
CC  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AY073600; AAL61263.1; -  
CC EMBL; AY317600; AAP70995.1; -  
CC InterPro: IPR000276; GPCR\_Rhodpsn.  
CC Pfam; PF00001; 7tm.1.1  
CC PRINTS; PR00237; GPCRHOOPS.  
CC PROSITE; PS00237; G-PROTEIN RECF\_F1\_1; 1.  
CC PROSITE; PS0262; G-PROTEIN RECF\_F1\_2; 1.  
CC G-protein coupled receptor; Transmembrane; Glycoprotein;  
KW Multigene family; Olfaction.  
KW DOMAIN 1 28 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 29 49 1 (POTENTIAL).  
FT TRANSMEM 50 57 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 58 78 2 (POTENTIAL).  
FT DOMAIN 79 102 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 103 123 3 (POTENTIAL).  
FT DOMAIN 124 136 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 137 157 4 (POTENTIAL).  
FT DOMAIN 158 199 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 200 220 5 (POTENTIAL).  
FT DOMAIN 221 240 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 241 261 6 (POTENTIAL).  
FT DOMAIN 262 274 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 275 295 7 (POTENTIAL).  
FT DOMAIN 296 314 CYTOPLASMIC (POTENTIAL).  
FT DISULFID 100 192 BY SIMILARITY.  
FT CARBOHYD 8 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 314 AA; 34716 MW; 6DBIA4BB76B8C795 CRC64;

Query Match 72.3%; Score 34; DB 1; Length 314;  
Best Local Similarity 77.8%; Pred. No. 20; Indels 0; Gaps 0;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TVSGNULTI 9  
|||||: ||  
Db 41 TVSGNLSTI 49

RESULT 12  
OR36 MOUSE STANDARD; PRT; 314 AA.  
AC Q8VFD1;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Putative olfactory receptor 204-21.  
GN MOR204-21.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21676863; PubMed=11802173;  
RA Zhang X., Firestein S.;  
RT "The olfactory receptor gene superfamily of the mouse.";  
RL Nat. Neurosci. 5:124-133(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21864068; PubMed=11875048;  
RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tonnes-Priddy L.,  
RA Trask B.J.;  
RT "Different evolutionary processes shaped the mouse and human  
olfactory receptor gene families.";  
RL Hum. Mol. Genet. 11:535-546(2002).  
RN [3]  
RP ERRATUM.  
RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tonnes-Priddy L.,  
RA Trask B.J.;  
RL Hum. Mol. Genet. 11:1683-1683(2002).  
CC -!- FUNCTION: Potential odorant receptor.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
CC  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AY073672; AAL61335.1; -  
CC EMBL; AY317577; AAP70975.1; -  
CC InterPro: IPR000276; GPCR\_Rhodpsn.  
CC Pfam; PF00001; 7tm.1.1  
CC PRINTS; PR00237; GPCRHOOPS.  
CC PROSITE; PS00237; G-PROTEIN RECF\_F1\_1; 1.  
CC PROSITE; PS0262; G-PROTEIN RECF\_F1\_2; 1.  
CC G-protein coupled receptor; Transmembrane; Glycoprotein;  
KW Multigene family; Olfaction.  
KW DOMAIN 1 28 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 29 49 1 (POTENTIAL).  
FT TRANSMEM 50 57 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 58 78 2 (POTENTIAL).  
FT TRANSMEM

```

FT DOMAIN 79 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 123 3 (POTENTIAL).
FT DOMAIN 124 136 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 137 157 4 (POTENTIAL).
FT DOMAIN 158 199 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 200 220 5 (POTENTIAL).
FT DOMAIN 221 240 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 241 261 6 (POTENTIAL).
FT DOMAIN 262 274 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 275 295 7 (POTENTIAL).
FT DOMAIN 296 314 CYTOPLASMIC (POTENTIAL).
FT DISULFID 100 192 BY SIMILARITY.
FT CARBOHYD 8 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 314 AA; 34839 MW; F19098FA1E3D3C7D CRC64;

Query Match 72.3%; Score 34; DB 1; Length 314;
Best Local Similarity 77.8%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TVSGNLTII 9
Db 41 TVSGNLTII 49

RESULT 13
OR39 MOUSE STANDARD; PRT; 314 AA.
AC Q8VEN6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative olfactory receptor 204-34.
GN MOR204-34.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21676863; PubMed=11875048;
RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tonnes-Priddy L.,
RA Trask B.J.;
RT "The olfactory receptor gene superfamily of the mouse.";
RL Hum. Mol. Genet. 11:535-546(2002).
RN [3]
RP ERRATUM.
RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tonnes-Priddy L.,
RA Trask B.J.;
RL Hum. Mol. Genet. 11:1683-1683(2002).
CC -!- FUNCTION: Potential odorant receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; AY073782; AAL61445.1;
CC EMBL; AY317618; AAP71008.1;
CC InterPro; IP0000276; GPCR_Rhodopsn.
CC Pfam; PF00001; 7tm.1; 1.
CC PRINTS; PR00237; GPCRHOPOPSN.

```

```

DR PROSITE; PS00237; G PROTEIN RECEPTOR F1.1; 1.
DR PROSITE; PS00262; G PROTEIN RECEPTOR F1.2; 1. Glycoprotein;
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family; Olfaction.
FT DOMAIN 1 28 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 29 49 1 (POTENTIAL).
FT DOMAIN 50 57 2 (POTENTIAL).
FT TRANSMEM 58 78 2 (POTENTIAL).
FT DOMAIN 79 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 123 3 (POTENTIAL).
FT DOMAIN 124 136 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 137 157 4 (POTENTIAL).
FT DOMAIN 158 199 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 200 220 5 (POTENTIAL).
FT DOMAIN 221 240 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 241 261 6 (POTENTIAL).
FT DOMAIN 262 274 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 275 295 7 (POTENTIAL).
FT DOMAIN 296 314 CYTOPLASMIC (POTENTIAL).
FT DISULFID 100 192 BY SIMILARITY.
FT CARBOHYD 8 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 314 AA; 34883 MW; F5B91DB9E649B0C7 CRC64;

Query Match 72.3%; Score 34; DB 1; Length 314;
Best Local Similarity 77.8%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TVSGNLTII 9
Db 41 TVSGNLTII 49

RESULT 14
OR40 MOUSE STANDARD; PRT; 314 AA.
ID OR40 MOUSE STANDARD; PRT; 314 AA.
AC Q8VEN5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative olfactory receptor 204-35.
GN MOR204-35.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21676863; PubMed=11802173;
RA Zhang X., Firestein S.;
RT "The olfactory receptor gene superfamily of the mouse.";
RL Nat. Neurosci. 5:124-133(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21864068; PubMed=11875048;
RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tonnes-Priddy L.,
RA Trask B.J.;
RT "Different evolutionary processes shaped the mouse and human
RT olfactory receptor gene families.";
RL Hum. Mol. Genet. 11:535-546(2002).
RN [3]
RP ERRATUM.
RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tonnes-Priddy L.,
RA Trask B.J.;
RL Hum. Mol. Genet. 11:1683-1683(2002).
CC -!- FUNCTION: Potential odorant receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; AY073782; AAL61445.1;
CC EMBL; AY317618; AAP71008.1;
CC InterPro; IP0000276; GPCR_Rhodopsn.
CC Pfam; PF00001; 7tm.1; 1.
CC PRINTS; PR00237; GPCRHOPOPSN.

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

CC -----  
 CC EMBL; AY073783; AAL61446.1; --  
 CC EMBL; AY317601; AAP70996.1; --  
 CC InterPro; IPR000276; GPCR\_Rhodopsn.  
 CC Pfam; PF00001; 7tm1.1; 1.  
 CC PRINTS; PR00237; GPCRHHODOPSN.  
 CC PROSITE; PS00237; G-PROTEIN RECF F1\_1; 1.  
 CC PROSITE; PS0262; G-PROTEIN RECF F1\_2; 1.  
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Multigene family; Olfaction.  
 CC -----

FT DOMAIN 1 28 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 29 49  
 FT DOMAIN 50 57 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 58 78  
 FT DOMAIN 79 102 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 103 123  
 FT DOMAIN 124 136 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 137 157  
 FT DOMAIN 158 199  
 FT TRANSMEM 200 220  
 FT DOMAIN 221 240  
 FT TRANSMEM 241 261  
 FT DOMAIN 262 274 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 275 295  
 FT DOMAIN 296 314 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 100 192 BY SIMILARITY.  
 FT CARBOHYD 8 8 N-LINKED (GLCNAC... (POTENTIAL).  
 SQ SEQUENCE 314 AA; 34953 MW; 99D8616859F6B54D CRC64;

Query Match 72.3%; Score 34; DB 1; Length 314;  
 Best Local Similarity 77.8%; Pred. No. 20;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 TVSGNLIIT 9  
 DB 41 TVSGNLSIT 49

## RESULT 15

OR22\_MOUSE STANDARD; PRT; 316 AA.  
 AC Q8VG13;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DE 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Putative olfactory receptor 204-7.  
 GN MOR204-7.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21676863; PubMed=11802173;  
 RA Zhang X., Firestein S.;  
 RT "The olfactory receptor gene superfamily of the mouse.";  
 RL Nat. Neurosci. 5:124-133(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21864068; PubMed=11875048;  
 RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tonnes-Priddy L.,  
 RA Trask B.J.;  
 RT "Different evolutionary processes shaped the mouse and human  
 RT olfactory receptor gene families.";  
 RL Hum. Mol. Genet. 11:535-546(2002).  
 RN [3]  
 RP ERRATUM  
 RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tonnes-Priddy L.,  
 RA Trask B.J.;  
 RL Hum. Mol. Genet. 11:1693-1693(2002).  
 CC -!- FUNCTION: Potential odorant receptor.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
 CC -----

DR EMBL; AY073352; AAL61015.1; --  
 DR EMBL; AY317615; AAP71005.1; --  
 DR InterPro; IPR000276; GPCR\_Rhodopsn.  
 DR Pfam; PF00001; 7tm1.1; 1.  
 DR PRINTS; PR00237; GPCRHHODOPSN.  
 DR PROSITE; PS00237; G-PROTEIN RECF F1\_1; FALSE\_NEG.  
 DR PROSITE; PS0262; G-PROTEIN RECF F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Multigene family; Olfaction.  
 CC -----

FT DOMAIN 1 28 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 29 49  
 FT DOMAIN 50 57 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 58 78  
 FT DOMAIN 79 102 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 103 123  
 FT DOMAIN 124 136 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 137 157  
 FT DOMAIN 158 199  
 FT TRANSMEM 200 220  
 FT DOMAIN 221 240  
 FT TRANSMEM 241 261  
 FT DOMAIN 262 274 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 275 297  
 FT DOMAIN 298 316 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 100 192 BY SIMILARITY.  
 FT CARBOHYD 8 8 N-LINKED (GLCNAC... (POTENTIAL).  
 SQ SEQUENCE 316 AA; 34824 MW; 27C4B39DCA981DDB CRC64;

Query Match 72.3%; Score 34; DB 1; Length 316;  
 Best Local Similarity 77.8%; Pred. No. 20;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 TVSGNLIIT 9  
 DB 41 TVSGNLSIT 49

Search completed: June 22, 2004, 08:51:58  
 Job time : 8 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OK protein - protein search, using sw model

Run on: June 22, 2004, 08:47:17 ; Search time 24 Seconds  
(without alignments)  
131.466 Million cell updates/sec

Title: US-09-529-206d-4\_COPY\_127\_136

Perfect score: 47

Sequence: 1 TVSGNLTIR 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_vodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriap:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	83.0	142	4 Q9NY13	Q9ny13 homo sapien
2	38	80.9	180	4 Q9Y479	Q9y479 homo sapien
3	37	78.7	250	16 Q8G743	Q8g743 bifidobacte
4	36	76.6	245	9 Q64327	Q64327 bacterioph
5	36	76.6	318	16 Q8EAQ4	Q8eaq4 shewanella
6	36	76.6	352	13 Q9YHY4	Q9yhy4 lampetra fl
7	36	76.6	450	16 Q9RJ91	Q9rj91 streptomyce
8	35	74.5	144	16 Q8A9J8	Q8a9j8 bacteroides
9	35	74.5	335	16 Q932F0	Q932f0 staphylococ
10	35	74.5	335	16 Q9YVW7	Q9yvw7 staphylococ
11	35	74.5	336	16 Q8C035	Q8c035 staphylococ
12	35	74.5	674	13 Q9PT83	Q9pt83 xenopus lae
13	35	74.5	674	13 Q9PT82	Q9pt82 xenopus lae
14	35	74.5	742	5 Q9VS00	Q9vs00 drosophila
15	35	74.5	742	5 Q8T4D0	Q8t4d0 drosophila
16	35	74.5	907	5 Q9NEG0	Q9neg0 drosophila

17	35	74.5	947	2 Q93FV1	Q93fvi bacteroides
18	35	74.5	1066	5 Q8MSR5	Q8msr5 drosophila
19	35	74.5	12615	5 Q8IRV7	Q8irv7 drosophila
20	35	74.5	4117	5 Q8IRV9	Q8irv9 drosophila
21	35	74.5	4179	5 Q9W4Y4	Q9w4y4 drosophila
22	35	74.5	4223	5 Q8MPN3	Q8mpn3 drosophila
23	35	74.5	4228	5 Q8IRV8	Q8irv8 drosophila
24	34	72.3	99	6 Q97675	Q97675 sus scrofa
25	34	72.3	305	11 Q8VGG6	Q8vvg6 mus musculu
26	34	72.3	314	11 Q7TRU6	Q7tru6 mus musculu
27	34	72.3	314	16 Q9CEZ4	Q9ceza lactococcus
28	34	72.3	316	11 Q8VFC3	Q8vfc3 mus musculu
29	34	72.3	325	16 Q8UJK3	Q8ujk3 agrobacteri
30	34	72.3	444	16 Q89GD9	Q89gd9 lactobacill
31	34	72.3	580	16 Q7VCN2	Q7vcn2 prochloroco
32	34	72.3	585	4 Q8WM52	Q8wm52 homo sapien
33	34	72.3	585	4 Q8NAX9	Q8nax9 homo sapien
34	34	72.3	633	5 Q8SZU5	Q8szu5 drosophila
35	34	72.3	880	16 Q8EW45	Q8ew45 mycoplasma
36	34	72.3	1038	5 Q8MOW9	Q8mgw9 drosophila
37	34	72.3	1266	16 Q8X9D5	Q8x9d5 escherichia
38	34	72.3	1686	5 Q86119	Q86119 dictyostell
39	34	72.3	1749	2 Q9KGY5	Q9kgv5 alteromonas
40	34	72.3	2347	5 Q8INH9	Q8inh9 drosophila
41	34	72.3	2451	5 Q9VG05	Q9vg05 drosophila
42	33	70.2	95	9 Q38398	Q38398 bacterioph
43	33	70.2	104	16 Q8DH09	Q8dh09 synchococc
44	33	70.2	125	6 Q77674	Q77674 oryctolagus
45	33	70.2	144	2 Q47426	Q47426 escherichia

#### ALIGNMENTS

##### RESULT 1

Q9NY13 ID Q9NY13 PRELIMINARY; PRT; 142 AA.  
AC Q9NY13;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DB Hypothetical protein (Fragment).  
GN LAGE-2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID=9606;  
RX [1]  
RP SEQUENCE FROM N.A.  
RA Lethe B.G.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ275978; CAB76945.1; -;  
KW Hypothetical protein.  
FT NON TER 1 1  
SQ SEQUENCE 142 AA; 13895 MW; 27EBE922AC4ACC7B CRC64;

Query Match 83.0%; Score 39; DB 4; Length 142;  
Best Local Similarity 88.9%; Pred. No. 5.3;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 9  
DB 101 TVSGNLTMT 109  
|||||||:  
|||||||:

##### RESULT 2

Q9Y479 ID Q9Y479 PRELIMINARY; PRT; 180 AA.  
AC Q9Y479;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE LAGE-1S protein (Cancer/testis antigen 2).

```

GN LAGE1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
DE [1]
GN
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=9932550; PubMed=10399963;
RA Arnould C.A., Van den Doel P.B., Heenskerk B., Schrier P.I.;
RT "Interleukin-2-induced, melanoma-specific T cells recognize CAMEL, an
RL unexpected translation product of LAGE-1.";
RN Int. J. Cancer 82:442-448(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Aradhy S., Bardaro T., Galgoczy P., Yamagata T., Esposito T.,
RA Patian H., Ciccodicola A., Kenwick S., Platzner M., D'Urso M.,
RA Nelson D.L.;
RT "Multiple pathogenic and benign genomic rearrangements occur at a 35-
RT kb duplication involving the NEMO and the LAGE2 genes.";
RL Hum. Mol. Genet. 0:0-0(2001).
RX Hum. Mol. Genet. 0:0-0(2001).
DR EMBL; AJ012834; CAA10194.1; -.
DR EMBL; AF277315; AAL27015.1; -.
SQ SEQUENCE 180 AA; 18236 MW; 9077FAF953543A25 CRC64;

Query Match 80.9%; Score 38; DB 4; Length 180;
Best Local Similarity 80.0%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10
Db 127 TVSGNLTIR 136

RESULT 3
Q8G743 PRELIMINARY; PRT; 250 AA.
AC Q8G743;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Possible cobyric acid synthase CobQ.
GN B0429.
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 2705;
RX MEDLINE=22294977; PubMed=12381787;
RA Scheil M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
RA Pessi G., Zwiars M.-C., Desiere F., Bork P., Delley M.,
RA Priddy R.D., Arigoni F.;
RT "The genome sequence of Bifidobacterium longum reflects its adaptation
RT to the human gastrointestinal tract.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
DR EMBL; AB014661; AAN24286.1; -.
KW Complete proteome.
SQ SEQUENCE 250 AA; 27653 MW; 7577954A609689CF CRC64;

Query Match 78.7%; Score 37; DB 16; Length 250;
Best Local Similarity 87.5%; Pred. No. 25;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SGNLTIR 10
Db 21 SGNLTIR 28

RESULT 4
O64327 PRELIMINARY; PRT; 245 AA.
ID O64327

```

```

AC O64327;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE G013.
GN
RN
RP SEQUENCE FROM N.A.
RC Bacteriophage N15.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=40631;
RN [1]
RP SEQUENCE FROM N.A.
RA Hendrix R.W., Ravin V.K., Casjens S.R., Ford M.E., Ravin N.V.,
RA Smirnov I.K.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF064539; AAC19050.1; -.
DR F01; T13099; T13099.
DR InterPro; IPR003343; Big 2.
DR InterPro; IPR008964; Invasin_intimin.
DR Pfam; PF02368; Big 2; 1.
DR SMART; SM00635; BID_2; 1.
SQ SEQUENCE 245 AA; 25571 MW; 2CC941997C103FD3 CRC64;

Query Match 76.6%; Score 36; DB 9; Length 245;
Best Local Similarity 77.8%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 9
Db 209 TVSGNLTIR 217

RESULT 5
Q8EAQ4 PRELIMINARY; PRT; 318 AA.
ID Q8EAQ4
AC Q8EAQ4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN S03842.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.P., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Imprim M., Lee K., Berry K., Lee C.,
RA Mueller J., Kouri H., Gill J., Utterback T.R., McDonald L.A., C.M.;
RA Meldrum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AB015817; AAN56819.1; -.
DR TIGR; S03842; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 318 AA; 34113 MW; 0A037165033B6FCF CRC64;

Query Match 76.6%; Score 36; DB 16; Length 318;
Best Local Similarity 87.5%; Pred. No. 52;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 8
Db 153 TVSGNLTIR 160

```

```

RT the 8 Mb Streptomyces coelicolor A3(2) chromosome." ;
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=A3(2) / M145;
RC MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S., Taylor K.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)." ;
RL Nature 417:141-147(2002).
DR EMBL; AL939106; CAS61191.1; -
DR HSP; P07986; IXP.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . . ; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001919; Bac celose-bind.
DR InterPro; IPR008965; Cellul bind.
DR InterPro; IPR001000; Glyco_Hydro_10.
DR Pfam; PF00553; CBM_2; 1.
DR Pfam; PF00331; Glyco_Hydro_10; 1.
DR PRINTS; PR00134; GLHYDRLASE10.
DR SMART; SM00637; CBD II; 1.
DR SMART; SM00633; Glyco_10; 1.
DR PROSITE; PS00591; GLYCOSYL HYDROL F10; 1.
DR GlycoSIDase; Hydrolase; Xylan degradation; Complete proteome.
KW GlycoSIDase; Hydrolase; Xylan degradation; Complete proteome.
SQ SEQUENCE 450 AA; 47090 MW; E2PF6A36EC29B6AE CRC64;

Query Match 76.6%; Score 36; DB 16; Length 450;
Best Local Similarity 60.0%; Pred. No. 75;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TVSGNLTIR 10
Db 405 TVSGNWTIVR 414

RESULT 8
Q8A9J8 PRELIMINARY; PRT; 144 AA.
AC Q8A9J8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Lipoprotein, function unknown.
GN B70817.
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=VPI-5482 / ATCC 29148;
RC MEDLINE=22550858; PubMed=12663928;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis." ;
RL Science 299:2074-2076(2003).
DR EMBL; AE016929; AAC075924.1; -
DR InterPro; IPR007298; N1PE.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR Pfam; PF04170; N1PE; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Complete proteome.
SQ SEQUENCE 144 AA; 15781 MW; AB28935887F6976B CRC64;

Query Match 74.5%; Score 35; DB 16; Length 144;

```

Best Local Similarity 60.0%; Pred. No. 36;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10  
|: |||||:  
Db 96 TLEGNLTIR 105

RESULT 9  
Q932FO PRELIMINARY; PRT; 335 AA.

AC Q932FO; (1)  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein SAV0664.  
GN SAV0664.  
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=158878;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21311952; PubMed=11418146;  
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,  
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;  
RT "Whole genome sequencing of methicillin-resistant Staphylococcus  
aureus."  
RL Lancet 357:1225-1240 (2001).  
DR EMBL; AP003360; BAB56826.1; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005315; F:inorganic phosphate transporter activity; IEA.  
DR GO; GO:0006817; P:phosphate transport; IEA.  
DR Pfam; PF01384; PHO4; 1.  
DR InterPro; IPR001204; Phos\_transporter.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 335 AA; 35604 MW; 9135CAB8343FD628 CRC64;

Query Match 74.5%; Score 35; DB 16; Length 335;  
Best Local Similarity 70.0%; Pred. No. 89;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10  
|: |||||:  
Db 245 TVGGINMKIR 254

RESULT 10  
Q99VV7 PRELIMINARY; PRT; 335 AA.

AC Q99VV7; (1)  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein SA0619 (MW0626 protein).  
GN SA0619 OR MW0626.  
OS Staphylococcus aureus (strain N315), and  
OS Staphylococcus aureus (strain MW2).  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=158879, 196620;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21311952; PubMed=11418146;  
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J., Ito T., Kanamori M.,  
RA Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y.,  
RA Takahashi N.K., Sawano T., Inoue R.-I., Kaito C., Sekimizu K.,  
RA Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,

Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,  
Ogasawara N., Hayashi H., Hiramatsu K.;  
RT "Whole genome sequencing of methicillin-resistant Staphylococcus  
aureus."  
RL Lancet 357:1225-1240 (2001).  
DR EMBL; AP003360; BAB56826.1; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005315; F:inorganic phosphate transporter activity; IEA.  
DR GO; GO:0006817; P:phosphate transport; IEA.  
DR InterPro; IPR001204; Phos\_transporter.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 335 AA; 35588 MW; 9121CBF8343FD628 CRC64;

Query Match 74.5%; Score 35; DB 16; Length 335;  
Best Local Similarity 70.0%; Pred. No. 89;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10  
|: |||||:  
Db 245 TVGGINMKIR 254

RESULT 11  
Q8CQ35 PRELIMINARY; PRT; 336 AA.

AC Q8CQ35; (1)  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Low-affinity inorganic phosphate transporter.  
GN S80432.  
OS Staphylococcus epidermidis.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=1282;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=ATCC 12228.  
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,  
RA Chen Z., Wen Y.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE016745; AAC04029.1; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005315; F:inorganic phosphate transporter activity; IEA.  
DR GO; GO:0006817; P:phosphate transport; IEA.  
DR InterPro; IPR001204; Phos\_transporter.  
DR Pfam; PF01384; PHO4; 1.  
KW Complete proteome.  
SQ SEQUENCE 336 AA; 35811 MW; 5918CA94BC9A3D99 CRC64;

Query Match 74.5%; Score 35; DB 16; Length 336;  
Best Local Similarity 70.0%; Pred. No. 89;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10  
|: |||||:  
Db 245 TVGGINMKIR 254

RESULT 12  
Q9PT83

ID Q9PT83 PRELIMINARY; PRT; 674 AA.  
 AC Q9PT83;  
 DT 01-MAY-2000 (TREMELrel. 13, Created)  
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
 DE Sodium phosphate cotransporter.  
 GN XNAPI-A.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Intestine;  
 RX MEDLINE=97248089; PubMed=9094212;  
 RA Ishizuya-Oka A., Stelow M.A., Ueda S., Shi Y.B.;  
 RT "Temporal and spatial expression of an intestinal Na+/PO4 3-  
 cotransporter correlates with epithelial transformation during thyroid  
 RT hormone-dependent frog metamorphosis.";  
 RL Dev. Genet. 20:53-66(1997).  
 DR EMBL; L78835; AAF21134.1;  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0015321; F:sodium-dependent phosphate transporter acti. .; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006817; P:phosphate transport; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR006209; EGF like.  
 DR InterPro; IPR000005; HTHARAC.  
 DR InterPro; IPR003841; Na/Pi cotranspt.  
 DR Pfam; PF02690; Na Pi cotrans; 2.  
 DR TIGRFAMs; TIGR01013; 2a58; 1.  
 DR PROSITE; PS00022; EGF 1; 1.  
 DR PROSITE; PS00041; HTH\_ARAC FAMILY 1; 1.  
 SQ SEQUENCE 674 AA; 73818 MW; E34B7809A346B5P4 CRC64;  
 Query Match 74.5%; Score 35; DB 13; Length 674;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 VSGNLTIR 10  
 Db 168 VSSNLTIVR 176  
 RESULT 13  
 Q9PT82 PRELIMINARY; PRT; 674 AA.  
 AC Q9PT82;  
 DT 01-MAY-2000 (TREMELrel. 13, Created)  
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
 DE Sodium phosphate cotransporter.  
 GN XNAPI-B.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Intestine;  
 RX MEDLINE=97248089; PubMed=9094212;  
 RA Ishizuya-Oka A., Stelow M.A., Ueda S., Shi Y.B.;  
 RT "Temporal and spatial expression of an intestinal Na+/PO4 3-  
 cotransporter correlates with epithelial transformation during thyroid  
 RT hormone-dependent frog metamorphosis.";  
 RL Dev. Genet. 20:53-66(1997).  
 DR EMBL; L78836; AAF21135.1;  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0015321; F:sodium-dependent phosphate transporter acti. .; IEA.

DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006817; P:phosphate transport; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR006209; EGF like.  
 DR InterPro; IPR000005; HTHARAC.  
 DR InterPro; IPR003841; Na/Pi cotranspt.  
 DR Pfam; PF02690; Na Pi cotrans; 2.  
 DR TIGRFAMs; TIGR01013; 2a58; 1.  
 DR PROSITE; PS00022; EGF 1; 1.  
 DR PROSITE; PS00041; HTH\_ARAC FAMILY 1; 1.  
 SQ SEQUENCE 674 AA; 73837 MW; 70CE7520C98B73F1 CRC64;  
 Query Match 74.5%; Score 35; DB 13; Length 674;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 VSGNLTIR 10  
 Db 168 VSSNLTIVR 176  
 RESULT 14  
 Q9VS00 PRELIMINARY; PRT; 742 AA.  
 ID Q9VS00  
 AC Q9VS00;  
 DT 01-MAY-2000 (TREMELrel. 13, Created)  
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
 DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)  
 DE CG10064 protein.  
 GN CG10064.  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OC NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkely;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Calnaker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Balwle R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwack C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AB003560; AAF50630.1; -  
 DR FlyBase; FBgn0035724; CG10064.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 10.  
 DR PRINTS; PR00320; GPROTEINRPT.  
 DR SMART; SM00320; WD40; 8.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; 1.  
 DR PROSITE; PS00082; WD\_REPEATS\_2; 4.  
 DR PROSITE; PS0294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 742 AA; 81994 MW; B8F3210064859141 CRC64;  
  
 Query Match 74.5%; Score 35; DB 5; Length 742;  
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
 Qy 2 VSGNLTIR 10  
 Db 154 ISGNALTIVR 162  
  
 RESULT 15  
 Q8T4D0 PRELIMINARY; PRT; 742 AA.  
 AC Q8T4D0;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 \*DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE AT03371P.  
 GN CG10064.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 \*OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 SN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Nuncio J., Pacleb J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Celisner S.,  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY089248; AAL89986.1; -  
 DR FlyBase; FBgn0035724; CG10064.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 10.  
 DR PRINTS; PR00320; GPROTEINRPT.  
 DR SMART; SM00320; WD40; 10.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; 1.  
 DR PROSITE; PS00082; WD\_REPEATS\_2; 4.  
 DR PROSITE; PS0294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 742 AA; 82012 MW; BC13210535849140 CRC64;  
  
 Query Match 74.5%; Score 35; DB 5; Length 742;  
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
 Qy 2 VSGNLTIR 10  
 Db 154 ISGNALTIVR 162  
  
 Search completed: June 22, 2004, 08:53:26  
 Job time : 27 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run On: June 22, 2004, 08:40:22 ; Search time 68.6667 Seconds  
(without alignments)  
82.295 Million cell updates/sec

Title: US-09-529-206d-4\_COPY\_43\_62

Perfect score: 107

Sequence: 1 RGRGAGARASPGGGAPR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001s:\*

5: Geneseq2002s:\*

6: Geneseq2003as:\*

7: Geneseq2003bs:\*

8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	20	2	AAY05979 Human can
2	107	100.0	180	2	AAW62584 Cancer as
3	107	100.0	180	2	AAW69665 Human NY-
4	107	100.0	180	2	AAY05965 Human can
5	107	100.0	180	3	AAV52430 Human tum
6	107	100.0	180	3	AAV70862 Human tum
7	107	100.0	180	3	AAW03154 Human oes
8	107	100.0	180	4	AAW69946 Human NY-
9	107	100.0	180	4	AAW67164 Amino aci
10	107	100.0	180	4	AAU01535 Human NY-
11	107	100.0	180	4	AAW07714 Human NY-
12	107	100.0	180	5	AAU84818 Human NYN
13	107	100.0	180	5	AAU11543 Human tum
14	107	100.0	180	5	ABR58672 Human can
15	107	100.0	180	6	ABR48210 Human bla
16	107	100.0	180	6	ABU56508 Lung can
17	107	100.0	180	6	ABU56694 Lung can
18	107	100.0	180	6	ABP74198 Human NY-
19	107	100.0	180	6	ABR83438 Human NY-
20	107	100.0	180	7	ADC09577 Human NY-
21	107	100.0	180	7	ADD35568 Human NY-
22	107	100.0	180	7	ADD25510 Binding d
23	107	100.0	397	4	AAE13122 NY-ESO-IC
24	102	95.3	30	5	AAU85105 Human NYN
25	102	95.3	3541	5	AAU85130 Human mel

26	101	94.4	180	6	ABU64816 Human NY-
27	101	94.4	180	7	ADD35564 Human NY-
28	99	92.5	135	6	ABR58673 Human can
29	99	92.5	135	6	ABR48211 Human bla
30	99	92.5	135	6	ABU56509 Lung can
31	99	92.5	135	6	ABU56695 Lung can
32	99	92.5	180	2	AAW69664 Human IAG
33	99	92.5	180	3	AAV70860 Human IAG
34	99	92.5	180	3	AAU84820 Human LAG
35	99	92.5	180	5	ABW78346 Amino aci
36	99	92.5	180	6	ABP74199 Human LAG
37	99	92.5	180	7	ADC09577 LAGB-1a p
38	99	92.5	180	7	ADC35103 Human bre
39	99	92.5	210	2	AAW69663 Human LAG
40	99	92.5	210	3	AAV70861 Human LAG
41	99	92.5	210	6	ADA83806 Human CTA
42	99	92.5	210	6	ABP74200 Human LAG
43	99	92.5	210	7	ADC09578 LAGB-1b p
44	94	87.9	30	5	AAU85121 Human LAG
45	85	79.4	30	5	AAU85104 Human NYN

ALIGNMENTS

RESULT 1  
AAY05979  
ID AAY05979 standard; peptide; 20 AA.  
XX  
AC AAY05979;  
XX  
DT 16-AUG-1999 (first entry)  
XX  
DE Human cancer antigen NY ESO-1/CAG-3 ORF1 cancer peptide.  
XX  
KW NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
KW vaccine.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
FN WO9918206-A2.  
XX  
PD 15-APR-1999.  
XX  
PF 21-SEP-1998; 98WO-US019609.  
XX  
PR 08-OCT-1997; 97US-0061428P.  
XX  
FA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
FI Wang RF, Rosenberg SA;  
XX  
DR MPI, 1999-277270/23.  
XX  
PT Cancer antigen NY ESO1/CAG-3.  
XX  
PS Claim 16; Page 64; 88pp; English.  
XX

The present sequence represents a cancer peptide that is based on amino acid residues 44-62 of human ESO-1/CAG-3 (or CAG-3) ORF1 (see AAY05965), a new and potent tumour antigen capable of eliciting an antigen specific immune response by T cells. Cancer peptides derived from CAG-3 ORF1, CAG-3 ORF2 (see AAY05966), portions of them and their variants (see AAY05967-87), are useful as cancer vaccines that protect against cancer. The invention provides: vectors and host cells (also useful as vaccines); a method of diagnosis of cancer or precancer; a transgenic animal; antisense oligonucleotides that inhibit expression of the cancer peptide

CC or tumour antigen; antibodies reacting with a CAG-3 cancer peptide,  
CC useful in diagnostic and detection assays; and methods for preventing or  
CC inhibiting cancer by administering a cancer peptide, with or without an  
CC HLA molecule. The cancer peptides form part of, or are derived from,  
CC cancers such as primary or metastatic melanoma, thymoma, lymphoma,  
CC sarcoma, lung cancer, liver cancer, leukaemia, uterine cancer, cervical  
CC cancer, bladder cancer, kidney cancer and adenocarcinomas such as breast,  
CC prostate, ovarian, pancreatic and thyroid cancers. Melanoma is treated by  
CC inducing cancer-specific T cells in vitro for subsequent return to a  
CC patient  
XX  
SQ Sequence 20 AA;

Query Match 100.0%; Score 107; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 7.3e-06; Mismatches 0; Indels 0; Gaps 0;  
Matches 20; Conservative 0;

QY 1 RGRGAGAAASGPGGAPR 20  
| | | | | | | | | | | | | | | | | | | | | |  
DB 1 RGRGAGAAASGPGGAPR 20

RESULT 2  
AAW62584  
ID AAW62584 standard; protein; 180 AA.  
XX  
AC AAW62584;  
XX  
DT 17-SEP-1998 (first entry)  
XX  
DE Cancer associated antigen NY-ESO-1.  
XX  
KW Cancer associated antigen; NY-ESO-1; regression; progression; onset;  
XX cancer; treatment; diagnosis.  
XX  
OS Homo sapiens.  
XX

Key Location/Qualifiers  
FT Misc-difference 7 /note= "potential myristoylation site"  
FT Misc-difference 9 /note= "potential myristoylation site"  
FT Misc-difference 11 /note= "potential phosphorylation site"  
FT Misc-difference 98 /note= "potential phosphorylation site"  
FT Misc-difference 134 /note= "potential phosphorylation site"  
FT Misc-difference 138 /note= "potential phosphorylation site"  
FT Misc-difference 138 /note= "potential phosphorylation site"  
XX  
PN WO9814464-A1.  
XX  
PD 09-APR-1998.  
XX  
PF 15-SEP-1997; 97WO-US016335.  
XX  
PR 03-OCT-1996; 96US-00725182.  
XX  
PA (LUDW-) LUDWIG INST CANCER RES.  
XX  
PI Chen Y, Scanlan M, Gure A, Old LJ, Jager B, Knuth A;  
XX Drijfhout JW;  
XX  
XX WPI; 1998-286417/25.  
XX  
XX N-PSDB; AAV38566.  
XX

XX New isolated cancer associated antigen - is used to develop products for  
XX the diagnosis and treatment of cancers and for monitoring cancer therapy.  
XX  
XX Claim 8; Fig 3; 49pp; English.  
XX  
XX The present sequence represents a cancer associated antigen. The clone

CC from which the DNA sequence is obtained is designated NY-ESO-1. The  
CC specification described a method for determining regression, progression  
CC of onset of a cancerous condition, comprising monitoring a sample from a  
CC patient with the cancerous condition for a parameter selected from NY-ESO  
CC -1 protein, a peptide derived from NY-ESO-1 protein and cytolytic T cells  
CC specific for the peptide and an MHC molecule with which it non-covalently  
CC complexes. Methods for the treatment of a cancerous condition are also  
CC described. The NY-ESO-1 protein and peptides derived from it can be used  
CC for diagnosis and treatment of cancers and to monitor the efficacy of a  
CC therapeutic regime  
XX  
SQ Sequence 180 AA;

Query Match 100.0%; Score 107; DB 2; Length 180;  
Best Local Similarity 100.0%; Pred. No. 5.6e-05;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGRGAGAAASGPGGAPR 20  
| | | | | | | | | | | | | | | | | | | | | |  
DB 43 RGRGAGAAASGPGGAPR 62

RESULT 3  
AAW69665  
ID AAW69665 standard; protein; 180 AA.  
XX  
AC AAW69665;  
XX  
DT 27-OCT-1998 (first entry)  
XX  
DE Human NY-ESO-1 protein sequence, formerly known as LL-1.2 clone.  
XX  
KW Human; LAGE-1; tumour associated protein; LL-1; diagnosis; tumour.  
XX  
OS Homo sapiens.  
XX  
PN WO9832855-A1.  
XX  
PD 30-JUL-1998.  
XX  
PF 27-JAN-1998; 98WO-US001445.  
XX  
PR 27-JAN-1997; 97US-00791495.  
XX  
PA (LUDW-) LUDWIG INST CANCER RES.  
XX  
PI Lethe B, Lucas S, De Smet C, Godelaine D, Boon-Falleur T;  
XX WPI; 1998-427951/36.  
XX  
XX N-PSDB; AAV50348.  
XX  
XX New isolated LAGE-1 tumour associated nucleic acids - used to develop  
XX products for the diagnosis and treatment of LAGE-1 associated disorders,  
XX particularly tumours.

XX Example 2; Page 57-58; 73pp; English.  
XX  
XX The present sequence represents human NY-ESO-1, formerly known as LL-1.2  
XX clone, which is used in an example from the present invention which  
XX describes LAGE-1 tumour associated protein (TAP). The present invention  
XX also describes: (1) a method for treating a subject with a disorder  
XX characterised by expression of a LAGE-1 nucleic acid molecule or an  
XX expression product, comprising administering to the subject autologous  
XX cytolytic T cells to ameliorate the disorder, where the cytolytic T cells  
XX are specific for complexes of an HLA molecule and a LAGE-1 TAP or an  
XX immunogenic fragment; (2) a method for treating a subject with a disorder  
XX characterised by expression of a LAGE-1 nucleic acid molecule or an  
XX expression product, comprising administering a LAGE-1 TAP or an  
XX immunogenic fragment to ameliorate the disorder; and (3) a method for  
XX selectively enriching a population of T cells with cytolytic T cells  
XX specific for a LAGE-1 TAP comprising contacting an isolated population of  
XX T cells with an agent presenting a complex of a LAGE TAP or an  
XX immunogenic fragment and a HLA presenting molecule to selectively enrich

CC the isolated population of T cells with the cytolytic T cells. The  
CC methods and products from the present invention can be used for the  
CC diagnosis and treatment of LAGE-1 associated disorders, particularly  
CC tumours  
XX  
SQ Sequence 180 AA;  
Query Match 100.0%; Score 107; DB 2; Length 180;  
Best Local Similarity 100.0%; Pred. No. 5.6e-05;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RGPAGAGAAASGPGGAPR 20  
DB 43 RGPAGAGAAASGPGGAPR 62  
RESULT 4  
AA05965  
ID AAY05965 standard; protein; 180 AA.  
XX AC AAY05965;  
XX DT 16-AUG-1999 (first entry)  
XX DE Human cancer antigen NY ESO-1/CAG-3 ORF1 protein.  
XX KW NY ESO-1/CAG-3 gene; CAG-3 gene; cancer antigen; human; leukaemia;  
KW non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer; metastasis;  
KW melanoma; adenocarcinoma; thymoma; colon cancer; uterine cancer;  
KW breast cancer; prostate cancer; ovarian cancer; cervical cancer;  
KW bladder cancer; kidney cancer; pancreatic cancer; liver cancer; sarcoma;  
KW tumour; diagnosis; immunotherapy; therapy; vaccine; ORF1.  
XX OS Homo sapiens.  
XX PN W09918206-A2.  
XX PD 15-APR-1999.  
XX PF 21-SEP-1998; 98WO-US019609.  
XX PR 08-OCT-1997; 97US-0061428P.  
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX PI Wang RF, Rosenberg SA;  
XX DR WPI; 1999-277270/23.  
XX DR N-PSDB; AAX58599.  
XX PT Cancer antigen NY ESO1/CAG-3.  
XX PS Claim 4; Fig 3A; 88pp; English.  
CC The present sequence represents the ORF1 protein encoded by open reading  
CC frame 1 of the human ESO-1/CAG-3 (or CAG-3) gene. CAG-3 is a new and  
CC potent tumour antigen capable of eliciting an antigen specific immune  
CC response by T cells. Cancer peptides comprising ORF1, ORF2 (see  
CC AAY05965), portions of these peptides and their variants (see AAY05965-  
CC 87), are useful as cancer vaccines that protect the recipient from  
CC development of cancer. The invention provides: vectors and host cells  
CC (also useful as vaccines); a method of diagnosis of cancer or precancer;  
CC a transgenic animal; antisense oligonucleotides that inhibit expression  
CC of the cancer peptide or tumour antigen; antibodies reacting with the CAG  
CC -3 cancer peptide, useful in diagnostic and detection assays, and methods  
CC for preventing or inhibiting cancer by administering a cancer peptide,  
CC with or without an HLA molecule. The cancer peptides form part of, or are  
CC derived from, cancers such as primary or metastatic melanoma, thymoma,  
CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine cancer,  
CC cervical cancer, bladder cancer, kidney cancer and adenocarcinomas such  
CC as breast, prostate, ovarian, pancreatic and thyroid cancers. Melanoma is  
CC treated by inducing cancer-specific T cells in vitro for subsequent  
CC return to a patient

XX SQ Sequence 180 AA;  
Query Match 100.0%; Score 107; DB 2; Length 180;  
Best Local Similarity 100.0%; Pred. No. 5.6e-05;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RGPAGAGAAASGPGGAPR 20  
DB 43 RGPAGAGAAASGPGGAPR 62  
RESULT 5  
AAY52430  
ID AAY52430 standard; protein; 180 AA.  
XX AC AAY52430;  
XX DT 15-FEB-2000 (first entry)  
XX DE Human tumour antigen NY-ESO-1.  
XX KW Cancer; tumour; antigen; MHC; major histocompatibility complex; T-cell;  
KW cytotoxic; helper; stimulation; proliferation; treatment; diagnosis;  
KW prevention; melanoma; breast cancer; ovarian cancer; prostate cancer;  
KW hepatoma; thyroid cancer; bladder cancer; lung cancer; lymphoma.  
XX OS Homo sapiens.  
XX SH Key Location/Qualifiers  
FT Peptide 44..53 /note= "Peptide presented by MHC Class I HLA-B\*7"  
FT Peptide 60..69 /note= "Peptide presented by MHC Class I HLA-B\*7"  
FT Peptide 60..68 /note= "Peptide presented by MHC Class I HLA-B\*7"  
FT Peptide 63..72 /note= "Peptide presented by MHC Class I HLA-B\*7"  
FT Peptide 79..88 /note= "Peptide presented by MHC Class I HLA-B\*7"  
FT Peptide 79..87 /note= "Peptide presented by MHC Class I HLA-B\*7 and HLA-B\*35"  
FT Peptide 82..91 /note= "Peptide presented by MHC Class I HLA-A\*1"  
FT Peptide 82..90 /note= "Peptide presented by MHC Class I HLA-A\*1"  
FT Peptide 83..91 /note= "Peptide presented by MHC Class I HLA-B\*44"  
FT Peptide 84..92 /note= "Peptide presented by MHC Class I HLA-B\*8  
FT Peptide 87..96 /note= "Peptide presented by MHC Class I HLA-A\*1"  
FT Peptide 88..96 /note= "Peptide presented by MHC Class I HLA-B\*44"  
FT Peptide 96..104 /note= "Peptide presented by MHC Class I HLA-B\*7"  
FT Peptide 100..108 /note= "Peptide presented by MHC Class I HLA-B\*44"  
FT Peptide 102..110 /note= "Peptide presented by MHC Class I HLA-B\*44"  
FT Peptide 107..116 /note= "Peptide presented by MHC Class I HLA-A\*24"  
FT Peptide 110..118 /note= "Peptide presented by MHC Class I HLA-B\*52"  
FT Peptide 113..122 /note= "Peptide presented by MHC Class I HLA-B\*7 and HLA-B\*52"  
FT Peptide 113..121 /note= "Peptide presented by MHC Class I HLA-B\*7"  
FT Peptide 115..124 /note= "Peptide presented by MHC Class I HLA-A\*3"

FT Peptide 118..126  
 FT /note= "Peptide presented by MHC Class I HLA-B\*35"  
 FT Peptide 124..133  
 FT /note= "Peptide presented by MHC Class I HLA-B\*52"  
 FT Peptide 125..133  
 FT /note= "Peptide presented by MHC Class I HLA-A\*24"  
 FT Peptide 138..147  
 FT /note= "Peptide presented by MHC Class I HLA-B\*8"  
 FT Peptide 139..147  
 FT /note= "Peptide presented by MHC Class I HLA-B\*7"  
 FT Peptide 145..153  
 FT /note= "Peptide presented by MHC Class I HLA-A\*24 and HLA-B\*52"  
 FT Peptide 153..162  
 FT /note= "Peptide presented by MHC Class I HLA-B\*52"  
 FT Peptide 154..163  
 FT /note= "Peptide presented by MHC Class I HLA-B\*52"  
 FT Peptide 154..162  
 FT /note= "Peptide presented by MHC Class I HLA-B\*52"  
 FT Peptide 156..167  
 FT /note= "Peptide (AAV52434) presented by MHC Class I HLA-A\*2"  
 FT Peptide 158..166  
 FT /note= "Peptide presented by MHC Class I HLA-A\*3"  
 FT Peptide 159..167  
 FT /note= "Peptide presented by MHC Class I HLA-A\*3"  
 FT Peptide 162..170  
 FT /note= "Peptide presented by MHC Class I HLA-B\*52"  
 XX WO9953938-A1.  
 XX 28-OCT-1999.  
 XX 24-MAR-1999; 99WO-US006875.  
 XX 17-APR-1998; 98US-00062422.  
 XX 02-OCT-1998; 98US-00165546.  
 XX (LUDW-) LUDWIG INST CANCER RES.  
 XX Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;  
 XX Gure A, Ritter G;  
 XX WPI; 2000-038483/03.  
 XX N-PSDB; AA238380.  
 XX Novel peptides which bind to MHC class I and MHC class II molecules,  
 XX useful for therapeutic and diagnostic purposes.  
 XX Claim 30; Fig 3; 49pp; English.  
 CC This sequence represents a human tumour antigen, NY-ESO-1, the cDNA  
 CC encoding which was isolated from an oesophagus squamous cell cancer cDNA  
 CC library. Tissue localisation studies revealed it to be expressed at high  
 CC levels in normal ovary and testis but not in normal colon, kidney, liver,  
 CC brain, oesophagus and skin. It was expressed in certain tumours and  
 CC tumour cell lines with some degree of frequency - these included melanoma  
 CC specimens and cell lines, and breast and bladder cancer specimens, with  
 CC expression in other tumour types being sporadic. Peptides derived from NY  
 CC -ESO-1 are bound by both MHC (major histocompatibility complex) Class I  
 CC and Class II molecules for presentation to T-cells. Peptides AAV52431-  
 CC Y52434 bind to Class I HLA-A2 molecules, thereby stimulating  
 CC proliferation of cytotoxic T-cells, while peptides AAV52435-Y52440 bind  
 CC to Class II HLA-DR3 molecules, stimulating helper T-cell proliferation.  
 CC The peptides derived from NY-ESO-1 may be used in methods and  
 CC compositions used for the treatment, diagnosis and prevention of cancers  
 CC (such as melanoma, breast cancer, prostate cancer, lung cancer, hepatoma,  
 CC ovarian cancer, thyroid cancer, bladder cancer, or lymphoma) and to  
 CC stimulate the proliferation of T cells  
 XX Sequence 180 AA;

Query Match 100.0%; Score 107; DB 3; Length 180;

Best Local Similarity 100.0%; Pred. No. 5.6e-05; Indels 0; Gaps 0;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGRGGAARASGPGGAPR 20  
 DB 43 RGRGGAARASGPGGAPR 62  
 RESULT 6  
 AAY70862  
 ID AAY70862 standard; protein; 180 AA.  
 XX AC AAY70862;  
 XX DT 31-JUL-2000 (first entry)  
 XX DE Human tumour antigen, NY-ESO-1 protein.  
 XX KW NY-ESO-1; CAMEL; CTL-recognised Antigen on MELANOMA; human; cancer; CTL;  
 XX KW cytotoxic T lymphocyte; tumour-associated antigen; LAGE-1; anticancer;  
 XX KW melanoma; immunotherapy; immune response.  
 XX OS Homo sapiens.  
 XX PN WO200023584-A1.  
 XX PD 27-APR-2000.  
 XX PF 15-OCT-1999; 99WO-EP007832.  
 XX PR 16-OCT-1998; 98EP-00119583.  
 XX PA (BOBH ) BOEHRINGER INGELHEIM INT GMBH.  
 XX PA (UYHO-) UNIV HOSPITAL LEIDEN.  
 XX PI Schrier PI, Aarnoudse CA, Heider K, Klade C;  
 XX WPI; 2000-339685/29.  
 XX DR N-PSDB; AAD00152.  
 XX PT Tumor-associated antigen useful for cancer immunotherapy is encoded by  
 XX the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA.  
 XX Example 3; Page 62-63; 73pp; English.  
 CC The present sequence is the human NY-ESO-1 protein, a tumour antigen,  
 CC identified by screening an esophagus carcinoma cDNA library. This protein  
 CC is derived from open reading frame (ORF)-1 that contain epitopes of  
 CC tumour specific T-cells. NY-ESO-1 is expressed in different tumour types,  
 CC but not in healthy tissues except in testis. It also shows homology with  
 CC the CAMEL (Cytotoxic T lymphocytes (CTL)-recognised Antigen on MELANOMA)  
 CC protein, a tumour-associated antigen. The tumour-associated antigen  
 CC displayed on melanoma cells is recognised by cytotoxic T lymphocytes.  
 CC This sequence has anticancer activity. CAMEL tumour antigen and  
 CC immunogenic peptides derived from it are useful for cancer immunotherapy.  
 CC They have the potential to induce an immune response, by eliciting a CTL  
 CC response. The DNA molecule is used for the construction of recombinant or  
 CC fusion proteins  
 XX Sequence 180 AA;  
 QY 1 RGRGGAARASGPGGAPR 20  
 DB 43 RGRGGAARASGPGGAPR 62  
 RESULT 7  
 AAB03154  
 ID AAB03154 standard; protein; 180 AA.

Query Match 100.0%; Score 107; DB 3; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-05;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGRGGAARASGPGGAPR 20  
 DB 43 RGRGGAARASGPGGAPR 62  
 RESULT 7  
 AAB03154  
 ID AAB03154 standard; protein; 180 AA.

XX	AC	AAB03154;	
XX	XX		
XX	DT	23-OCT-2000	(first entry)
XX	XX		
XX	DE	Human oesophageal cancer-associated antigen NY-ESO-1.	
XX	XX		
KW	XX	Oesophageal cancer associated antigen; NY-ESO-1; human; immunogen;	
KW	XX	oesophageal carcinoma; melanoma; ovary; testis; transmembrane domain;	
KW	XX	antibody; diagnostic marker; drug delivery target.	
XX	OS	Homo sapiens.	
XX	XX		
XX	PH	Key	Location/Qualifiers
XX	FT	Modified-site	7
XX	FT		/note= "Potential N-myristoylation site"
XX	FT	Modified-site	9
XX	FT		/note= "Potential N-myristoylation site"
XX	FT	Modified-site	11
XX	FT		/note= "Potential O-phosphorylation site"
XX	FT	Modified-site	98
XX	FT		/note= "Potential O-phosphorylation site"
XX	FT	Modified-site	134
XX	FT		/note= "Potential O-phosphorylation site"
XX	FT	Modified-site	138
XX	FT		/note= "Potential O-phosphorylation site"
XX	FT	Domain	152..172
XX	FT		/note= "Potential transmembrane domain"
XX	PN	US6069233-A.	
XX	XX		
XX	PD	30-MAY-2000.	
XX	XX		
XX	PF	26-JAN-1998;	98US-00013150.
XX	XX		
XX	PR	03-OCT-1996;	96US-00725381.
XX	XX	{SLOK } SLOAN KETTERING INST CANCER RES.	
XX	PA	{CORR } CORNELL RES FOUND INC.	
XX	PA	{LUDW-} LUDWIG INST CANCER RES.	
XX	XX		
XX	PI	Scanlan M, Gure AO, Chen Y, Tureci O, Sahin U, Pfreundschuh M;	
XX	PI	Old Lu;	
XX	XX		
XX	DR	WPI; 2000-410880/35.	
XX	DR	N-PSDB; AAA61483.	
XX	XX		
XX	PT	New isolated esophageal cancer-associated antigen useful as markers for	
XX	PT	producing antibodies and as targets for identifying abnormal conditions,	
XX	PT	e.g. infections and cancer.	
XX	XX		
XX	PS	Example 5; Fig 3; 9pp; English.	
XX	XX		
XX	CC	This sequence represents a human oesophageal cancer-associated antigen,	
XX	CC	NY-ESO-1. The cDNA encoding this sequence was isolated from a cDNA	
XX	CC	library prepared from a specimen of well-to-moderately differentiated	
XX	CC	squamous cell cancer of the oesophagus. Expression analysis demonstrated	
XX	CC	that NY-ESO-1 was expressed in oesophageal carcinoma, certain melanoma	
XX	CC	cell lines and in normal ovary and testis tissue, but not in normal	
XX	CC	colon, kidney, liver or brain tissue. Analysis of the amino acid sequence	
XX	CC	of the protein indicates that the protein has a transmembrane domain,	
XX	CC	several N-myristoylation sites and O-phosphorylation sites and that it	
XX	CC	contains antigenic sequences in the N-terminal half of the protein. The	
XX	CC	antigen is useful as an immunogen when combined with an adjuvant. In both	
XX	CC	precursor and post- translationally modified forms, and may be used to	
XX	CC	generate anti-NY-ESO-1 antibodies. It can also be used as a diagnostic	
XX	CC	marker for oesophageal cancer, and can be utilised as a marker for the	
XX	CC	targetted delivery of therapeutic agents to oesophageal cancer cells. It	
XX	CC	can also be used to generate diagnostic or therapeutic agents	
XX	SQ	Sequence 180 AA:	

```

Best Local Similarity 100.0%; Pred. No. 5.6e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPRGAGARASPGGGAPR 20
DB 43 RPRGAGARASPGGGAPR 62

RESULT 8
AAB69946
ID AAB69946 standard; protein; 180 AA.
XX AAB69946;
XX
XX 27-APR-2001 (first entry)
XX Human NY-ESO-1 protein.
XX Human; NY-ESO-1; HLA; human leukocyte antigen; CTL; cytotoxic T cell;
XX HLA-A2; HLA-DR53; melanoma; adenocarcinoma; bladder carcinoma;
XX non-small cell lung carcinoma; tumour status determination.
XX
XX Homo sapiens.
XX WO200107917-A1.
XX
XX 01-FEB-2001.
XX
XX 14-JUL-2000; 2000WO-US019220.
XX
XX 23-JUL-1999; 99US-00359503.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX (SLOK) SLOAN KETTERING INST CANCER RES.
XX (CORR) CORNELL RES FOUND INC.
XX
XX Jager E, Stockert E, Old LJ, Knuth A, Chen Y, Scanlan M;
XX WPI; 2001-182822/18.
XX N-PSDB; AAP58634.
XX
XX Method useful for determining the status (e.g. progression, regression or
XX stability of the disease) of a cancerous condition, involves determining
XX the levels of NY-ESO-1 specific antibodies in a sample taken from a
XX patient.
XX
XX Example 5; Fig 3; 50pp; English.
XX
XX The present sequence is human NY-ESO-1 protein. It is provided in a
XX specification relating to a method for determining the status of a
XX cancerous condition in a patient with a tumour that expresses NY-ESO-1.
XX The method comprises assaying a sample taken from the patient for
XX antibodies that specifically bind to the NY-ESO-1 and comparing the value
XX obtained to a prior value obtained from assay of a prior sample taken
XX from the patient. Any difference between the values is indicative of a
XX change in status of the cancerous condition. The method is useful for
XX determining whether a cancerous condition is progressing, regressing or
XX remaining stable, in particular in patients receiving treatment for a
XX melanoma, adenocarcinoma, non-small cell lung carcinoma or bladder
XX carcinoma
XX
XX Sequence 180 AA;

Query Match 100.0%; Score 107; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 5.6e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPRGAGARASPGGGAPR 20
DB 43 RPRGAGARASPGGGAPR 62

```

## RESULT 9

Query Match  
100.0%; Score 107; DB 3; Length 180;

AAG67164  
ID AAG67164 standard; protein; 180 AA.  
XX  
AC AAG67164;  
XX  
DT 13-NOV-2001 (first entry)  
XX  
DE Amino acid sequence of cancer testis tumour antigen NY-ESO-1.  
XX  
KW Cancer testis tumour antigen; NY-ESO-1; LAGE-2; human leukocyte antigen;  
KW HLA; HLA binding peptide; major histocompatibility complex; MHC; tumour;  
KW cancer; testis tumour.  
XX  
OS Homo sapiens.  
XX  
PN WO200162917-A1.  
XX  
PD 30-AUG-2001.  
XX  
PF 22-JAN-2001; 2001WO-US002126.  
XX  
PR 22-FEB-2000; 2000US-00510635.  
XX  
PA (LUDW-) LUDWIG INST CANCER RES.  
XX  
PI Lethe B, Boon-Falleur T;  
XX  
DR WPI; 2001-550091/61.  
DR N-PSDB; AAS75118.  
XX  
PT Genomic sequences of tumor associated antigen NY-ESO-1 (LAGE-2) useful  
for diagnosing testicular tumors.  
XX  
PS Example 5; Fig 3; 50pp; English.  
XX  
CC The present sequence represents cancer testis tumour antigen NY-ESO-1  
(also called LAGE-2). NY-ESO-1 is a molecule that is processed to at  
least one human leukocyte antigen (HLA) binding peptide, which binds to  
Class I and Class II major histocompatibility complex (MHC). NY-ESO-1 is  
expressed in tumour mRNA and in testis, but not normal colon, kidney,  
liver or brain tissue. The presence or level of expression of NY-ESO-1  
may be assayed for the diagnosis of cancer, especially testis tumours  
XX  
SQ Sequence 180 AA;  
Query Match 100.0%; Score 107; DB 4; Length 180;  
Best Local Similarity 100.0%; Pred. No. 5.6e-05;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RGPRGAGAAASGPGGGAPR 20  
|||||  
DB 43 RGPRGAGAAASGPGGGAPR 62  
RESULT 10  
AAU01535  
ID AAU01535 standard; protein; 180 AA.  
XX  
AC AAU01535;  
XX  
DT 18-JUL-2001 (first entry)  
XX  
DE Human NY-ESO-1 tumour rejection antigen precursor protein.  
XX  
KW NY-ESO-1; human; tumour rejection antigen precursor; SSX-2; MHC Class II;  
KW major histocompatibility complex; helper T cell; HLA-DR; cancer;  
KW human leukocyte antigen-determining region; disease progression;  
KW disease regression; disease onset; body tissue; body fluid; enzyme label;  
KW radioactive label; monoclonal antibody.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers

Modified-site 7 /note= "Myristoylated"  
PT  
PT Modified-site 9 /note= "Myristoylated"  
PT  
PT Modified-site 11 /note= "Myristoylated"  
PT  
PT Modified-site 98 /note= "Phosphorylated"  
PT  
PT Modified-site 134 /note= "Phosphorylated"  
PT  
PT Modified-site 138 /note= "Phosphorylated"  
PT  
XX WO200123560-A2.  
XX  
XX PN  
XX PD 05-APR-2001.  
XX  
XX PF 26-SEP-2000; 2000WO-US026411.  
XX  
XX PR 29-SEP-1999; 99US-00408036.  
XX  
XX PA (LUDW-) LUDWIG INST CANCER RES.  
XX  
XX PI Tureci O, Sahin U, Pfreundschuh M;  
XX  
XX DR WPI; 2001-266156/27.  
XX  
XX DR N-PSDB; AAS02254.  
XX  
PT Polypeptides binding to major histocompatibility complex class II human  
leukocyte antigen-determining region molecule having amino acid sequence  
found in tumor rejection antigen precursor used for stimulating  
proliferation of helper T cells.  
XX  
PS Claim 4; Fig 3; 62pp; English.  
XX  
CC The sequence represents a human NY-ESO-1 tumour rejection antigen  
precursor. NY-ESO-1 and SSX-2 polypeptides, or fragments of, bind to  
major histocompatibility complex (MHC) Class II molecules such as human  
leukocyte antigen-determining region (HLA-DR) molecules and stimulate  
proliferation of helper T cells. The peptides can be administered to an  
HLA-DR positive subject in order to stimulate the helper T cells. An MHC  
Class II HLA-DR-NY-ESO-1/SSX-2 complex expressed on the surface of a cell  
or present in free form is useful for this stimulation. The nucleic acid  
is useful for screening for a cancerous condition, which involves  
contacting a subject sample to a cell line transfected with the  
immunoreactive cell (helper T cell), where interaction is indicative of  
cancer. In addition, a sample from a patient (for example, a body fluid  
or tissue) can be monitored for the amount of the complex present in the  
bloodstream. This is useful for determining regression, progression or  
onset of a cancerous condition. The method involves contacting the sample  
with a radioactive labelled or enzyme labelled monoclonal antibody which  
specifically binds with the complex  
XX  
SQ Sequence 180 AA;

Query Match 100.0%; Score 107; DB 4; Length 180;  
Best Local Similarity 100.0%; Pred. No. 5.6e-05;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RGPRGAGAAASGPGGGAPR 20  
|||||  
DB 43 RGPRGAGAAASGPGGGAPR 62  
RESULT 11  
AAE07714  
ID AAE07714 standard; protein; 180 AA.  
XX  
XX AC AAE07714;  
XX  
XX DT 06-NOV-2001 (first entry)  
XX  
XX DE Human NY ESO-1 protein.

```

XX Human; major histocompatibility complex; MHC; vaccine; metastasis;
KW class II restricted T cell epitope; MHC-II epitope; cancer antigen;
KW NY ESO-1 protein; CD4+ T lymphocyte; human leucocyte antigen; HLA;
KW tumour-specific humoral-mediated immunity; cancer; cytostatic;
KW immunotherapy.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX Misc-difference 45..47
XX /note= "Encoded by CCGGGGGC"
XX
XX WO200155393-A2.
XX
XX 02-AUG-2001.
XX
XX 26-JAN-2001; 2001WO-US002765.
XX
XX 28-JAN-2000; 2000US-0179004P.
XX 29-SEP-2000; 2000US-0237107P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Wang R, Rosenberg SA, Zeng G;
XX
XX WPI; 2001-496851/54.
XX N-ESDB; AAD14179, AAD14180.
XX
XX New NY-ESO cancer peptide or MHC class II restricted T cell epitopes,
XX useful as immunogen and vaccine for inhibiting cancer in a mammal or as
XX protection from metastasis.
XX
XX Example 1; Fig 1; 134pp; English.
XX
XX The invention relates to the identification and isolation of major
XX histocompatibility (MHC) class II restricted T cell epitope (MHC-II
XX epitope) derived from the cancer antigen, NY ESO-1. The MHC-II epitopes
XX from NY ESO-1 are recognised by CD4+ T lymphocytes in an human leucocyte
XX antigen (HLA) class II restricted manner, in particular HLA-DR or HLA-DP
XX restricted. The products of the gene are promising candidates for
XX immunotherapeutic strategies for the prevention, treatment and diagnosis
XX of patients with cancer. The cancer epitopes are useful as immunogen and
XX vaccine to inhibit or to prevent cancer in a mammal by eliciting CD4+ T
XX lymphocytes resulting in protection of the recipient from development of
XX cancer and protection from metastasis, or by inhibiting the growth of
XX cells expressing the NY-ESO-1 gene product. The cancer peptides are also
XX useful as diagnostic agent to detect the presence of cancer, to enhance
XX the generation of antibody and/or CD8+ T cell responses against any given
XX target antigen and/or hapten and to induce tumour-specific humoral-
XX mediated immunity against cancer. The present sequence is human NY ESO-1
XX protein
XX
XX SQ Sequence 180 AA;
XX
XX Query Match 100.0%; Score 107; DB 4; Length 180;
XX Best Local Similarity 100.0%; Pred. No. 5.6e-05;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 RCPRGAGAAASGPGGAPR 20
XX |||||
XX Db 43 RCPRGAGAAASGPGGAPR 62
XX
XX RESULT 12
XX AAU84818
XX ID AAU84818 standard; protein; 180 AA.
XX
XX AAU84818;
XX
XX 08-MAY-2002 (first entry)
XX
XX Human NYESO1a consensus sequence.
XX
XX DE

```

```

XX Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW viral infection; human immunodeficiency virus; melanoma;
KW bacterial infection; Salmonella; Legionella; parasitic infection;
KW Trypanosoma; Toxoplasma; Giardia.
XX
XX Homo sapiens.
OS
XX
XX WO200190197-A1.
XX
XX 29-NOV-2001.
XX
XX 25-MAY-2001; 2001WO-AU000622.
XX
XX 26-MAY-2000; 2000AU-00007761.
XX
XX (AUSU ) UNIV AUSTRALIAN NAT.
XX
XX Thomson SA, Ramshaw IA;
XX
XX WPI; 2002-147575/19.
XX
XX New synthetic polypeptides having several different segments of at least
XX one parent polypeptide linked together differently compared to the
XX linkage in the parent polypeptide, for inducing immune response against a
XX pathogen or cancer.
XX
XX Example 3; Fig 27; 364pp; English.
XX
XX The invention relates to a new synthetic polypeptide (I) comprising
XX several different segments of at least one parent polypeptide linked
XX together in a different relationship relative to their linkage in the
XX parent polypeptide to impede, abrogate or otherwise alter at least one
XX function associated with the parent polypeptide and for inducing an
XX immune response against a pathogen or cancer. Also included are a
XX synthetic polynucleotide encoding and a computer system for designing the
XX synthetic polypeptides. The synthetic polypeptides and polynucleotides
XX are referred to as a Savine. The synthetic polypeptide is useful for
XX modulating immune responses preferably directed against a pathogen or a
XX cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
XX and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
XX oesophagus, brain, testicle, uterus), as potentiating agents.
XX Compositions comprising the polypeptide may be used in the treatment or
XX prophylaxis against viral (such as infections caused by HIV (human
XX immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
XX virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
XX (e.g., infections caused by Neisseria, Meningococcus, Haemophilus,
XX Salmonella, Streptococcus, Legionella and Mycobacterium or parasitic
XX (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
XX Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
XX a consensus sequence for a parent protein used to design a savine of the
XX invention
XX
XX SQ Sequence 180 AA;
XX
XX Query Match 100.0%; Score 107; DB 5; Length 180;
XX Best Local Similarity 100.0%; Pred. No. 5.6e-05;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 RCPRGAGAAASGPGGAPR 20
XX |||||
XX Db 43 RCPRGAGAAASGPGGAPR 62
XX
XX RESULT 13
XX AAU11543
XX ID AAU11543 standard; protein; 180 AA.
XX
XX AAU11543;
XX
XX 12-MAR-2002 (first entry)
XX
XX Human tumour associated antigen NY-ESO.
XX
XX DE

```

XX Human; tumour associated antigen; NY-ESO; human leukocyte antigen;  
 KW major histocompatibility complex; MHC; HLA-A2.2; vaccine; cancer; HIV;  
 KW human immunodeficiency virus infection; cytostatic; virucide;  
 KW housekeeping epitope; adoptive immunotherapy; neoplastic disease;  
 KW viral disease; hepatitis virus; papilloma virus; tumour; leukaemia;  
 KW lymphoma; breast cancer; prostate cancer; lung cancer;  
 KW parasitic infection; Chlamydia; Trypanosoma; Toxoplasma.  
 XX Homo sapiens.  
 OS  
 XX WO200182963-A2.  
 PN  
 XX 08-NOV-2001.  
 PD  
 XX  
 XX 27-APR-2001; 2001WO-US013806.  
 PF  
 XX 28-APR-2000; 2000US-00560465.  
 PR 28-APR-2000; 2000US-00561074.  
 PR 28-APR-2000; 2000US-00561571.  
 PR 28-APR-2000; 2000US-00561572.  
 XX (CTLI-) CTL IMMUNOTHERAPIES CORP.  
 PA  
 XX Simard JUL, Diamond DC, Lei X;  
 PI  
 XX WPI; 2002-0566492/09.  
 DR  
 XX Novel vaccine useful for treating neoplastic and viral diseases,  
 PF comprises a first housekeeping epitope derived from a first antigen  
 PT associated with a first target cell.  
 PT  
 XX Example 23; Fig 17; 131pp; English.  
 PS  
 XX The invention relates to a vaccine comprising a first housekeeping  
 XX epitope derived from a first antigen associated with a first target cell.  
 CC Also included are an isolated T cell expressing a T cell receptor  
 CC specific for a major histocompatibility complex (MHC)-peptide complex  
 CC comprising a first housekeeping epitope which is derived from a first  
 CC antigen associated with a first target cell, selecting an epitope (or  
 CC peptide sequence) from a population of peptide fragments of an antigen  
 CC associated with a target in a host, where the fragments have a known or  
 CC predicted affinity for a MHC receptor peptide binding cleft of the host,  
 CC where the epitope selected corresponds to a product of proteolytic  
 CC cleavage of the antigen in a cell of the host and a nucleic acid  
 CC construct comprising a first coding region, where the first coding region  
 CC comprises a first sequence encoding at least a first polypeptide, where  
 CC the first polypeptide comprises a first housekeeping epitope derived from  
 CC a first antigen associated with a first target cell; The epitopes,  
 CC peptides, vaccines and nucleic acids are useful in the manufacture of a  
 CC medicament for use in adoptive immunotherapy and for prevention and  
 CC treatment of neoplastic and viral diseases (e.g. human immunodeficiency  
 CC virus, HIV, infection, hepatitis virus and papilloma virus), cancers  
 CC (e.g. tumours, leukaemia, lymphoma, breast cancer, prostate cancer and  
 CC lung cancer), infection of cells by intracellular parasites (e.g.  
 CC Chlamydia, Trypanosoma and Toxoplasma) and many other examples given in  
 CC the specification. The invention permits the vaccine designer to ignore  
 CC peptides that, despite predicted high binding affinity for MHC, will  
 CC never be useful because they cannot be presented by target cells. The  
 CC invention provides a major advance in vaccine design, one that combines  
 CC the power of antigen sequence analysis with the fundamental realities of  
 CC immunology. The invention allows for the simple and effective selection  
 CC of meaningful epitopes for creation of MHC class I or class II vaccines  
 CC using any polypeptide sequence corresponding to a desired target. The  
 CC present polypeptide is an HLA-A2.1 (human leukocyte antigen) presenting  
 CC target cell protein from which epitopes of the invention may be derived,  
 CC NY-ESO (a tumour associated antigen)  
 XX  
 SQ Sequence 180 AA;

Query Match 100.0%; Score 107; DB 5; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-05;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGPGRGAGARASGPGGAPR 20  
 Db |||||  
 43 RGPGRGAGARASGPGGAPR 62  
 RESULT 14  
 ABR58672  
 ID ABR58672 standard; protein; 180 AA.  
 XX  
 AC ABR58672;  
 XX  
 DT 09-JUL-2003 (first entry)  
 XX  
 DE Human cancer related protein SEQ ID NO:329.  
 XX  
 KW Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;  
 KW heart disease; atherosclerosis; endometriosis.  
 XX  
 OS Homo sapiens.  
 XX WO2003025138-A2.  
 PN  
 XX 27-MAR-2003.  
 PD  
 XX 17-SEP-2002; 2002WO-US029560.  
 PF  
 XX 17-SEP-2001; 2001US-0323469P.  
 PR 20-SEP-2001; 2001US-0323887P.  
 PR 13-NOV-2001; 2001US-0350666P.  
 PR 08-FEB-2002; 2002US-0355145P.  
 PR 08-FEB-2002; 2002US-0355257P.  
 PR 12-APR-2002; 2002US-0372246P.  
 XX  
 XX (EOSB-) EOS BIOTECHNOLOGY INC.  
 PA  
 XX Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;  
 PI Zlotnik A;  
 XX WPI; 2003-354600/33.  
 DR N-PSDB; ACC72823.  
 XX  
 XX New genes that are up-regulated or down-regulated in cancers, useful as  
 PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as  
 PT therapeutic targets for screening drugs for treating these diseases.  
 PT  
 XX Claim 12; Page 757-758; 767pp; English.  
 PS  
 XX The present invention describes an isolated nucleic acid molecule, which  
 CC comprises the sequence of any of the genes that are up-regulated or down-  
 CC regulated in specific cancers (e.g. about 1031 genes up-regulated in  
 CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer  
 CC related gene nucleotide sequences which encode the proteins given in  
 CC ABR58521 to ABR58709. Also described: (1) determining the presence or  
 CC absence of a pathological cell in a patient; (2) an expression vector  
 CC comprising a nucleic acid molecule described above; (3) a host cell  
 CC comprising the vector; (4) an isolated polypeptide, which is encoded by  
 CC the nucleic acid; (5) an antibody that specifically binds the polypeptide  
 CC of (4); (6) specifically targeting a compound to a pathological cell in a  
 CC patient by administering to the patient the antibody above; and (7) a  
 CC drug screening assay. The nucleic acid is useful as diagnostic markers or  
 CC therapeutic targets. In particular, the nucleic acid is useful for  
 CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,  
 CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,  
 CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,  
 CC atherosclerosis and endometriosis. The nucleic acid is also useful in  
 CC drug screening, particularly for identifying agents for treating these  
 CC pathologies  
 XX  
 SQ Sequence 180 AA;

Query Match 100.0%; Score 107; DB 6; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-05;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2004, 08:49:38 ; Search time 22 Seconds  
(without alignments)  
46.933 Million cell updates/sec

Title: US-09-529-206D-4\_COPY\_43\_62

Perfect score: 107

Sequence: 1 RGRGAGAAASGGGAPR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:\*
- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*
  - 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*
  - 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*
  - 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*
  - 5: /cgn2\_6/ptodata/2/iaa/ECTUS\_COMB.pep:\*
  - 6: /cgn2\_6/ptodata/2/iaa/backfiles.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	107	100.0	180	2	US-08-791-495-9
2	107	100.0	180	4	US-09-392-714-25
3	101	94.4	180	3	US-08-937-263B-8
4	101	94.4	180	4	US-09-751-798-8
5	99	92.5	180	2	US-08-791-495-7
6	99	92.5	210	2	US-08-791-495-5
7	63	58.9	860	4	US-09-252-991A-25681
8	60	56.1	160	4	US-09-252-991A-27091
9	60	56.1	928	1	US-08-442-248-2
10	60	56.1	928	1	US-08-440-815-2
11	60	56.1	928	3	US-08-486-449-2
12	60	56.1	928	4	US-08-578-684-2
13	60	56.1	1005	2	US-08-469-537A-103
14	59	55.1	210	4	US-09-252-991A-24923
15	59	55.1	330	1	US-08-118-270-21
16	59	55.1	330	5	PCR-US93-08528-21
17	59	55.1	877	2	US-08-673-789-2
18	58	54.2	456	4	US-09-252-991A-17335
19	58	54.2	700	4	US-09-252-991A-28344
20	56	52.3	421	4	US-09-252-991A-32326
21	56	52.3	432	4	US-09-252-991A-23622
22	55	51.4	191	4	US-09-252-991A-23951
23	55	51.4	289	4	US-09-252-991A-25496
24	55	51.4	343	4	US-09-252-991A-23398
25	54.5	50.9	501	4	US-09-252-991A-32473
26	54	50.5	140	3	US-09-220-528-5
27	54	50.5	140	4	US-09-347-613C-5

Sequence 10, Appl  
Sequence 12, Appl  
Sequence 89, Appl  
Sequence 40, Appl  
Sequence 18570, A  
Sequence 26, Appl  
Sequence 9, Appl  
Sequence 35, Appl  
Sequence 32, Appl  
Sequence 4, Appl  
Sequence 27073, A  
Sequence 24623, A  
Sequence 24722, A  
Sequence 59, Appl  
Sequence 2, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 60, Appl

US-09-347-613C-10  
US-09-220-528-12  
US-09-220-528-89  
US-09-220-528-40  
US-09-252-991A-18570  
US-09-220-528-26  
US-09-347-613C-9  
US-09-347-613C-35  
US-09-220-528-32  
US-09-347-613C-4  
US-09-252-991A-27073  
US-09-252-991A-24623  
US-09-252-991A-24722  
US-08-556-9788-59  
US-09-247-806-2  
US-08-425-069-4  
US-08-317-844B-4  
US-08-556-9788-60

#### ALIGNMENTS

RESULT 1  
US-08-791-495-9  
; Sequence 9, Application US/08791495  
; Patent No. 581519  
; GENERAL INFORMATION:  
; APPLICANT: Leth, Bernard  
; APPLICANT: Lucas, Sophie  
; APPLICANT: De Smet, Charles  
; APPLICANT: Godelaine, Danielle  
; APPLICANT: Boon-Palleur, Thierry  
; TITLE OF INVENTION: LI-1 TUMOR SPECIFIC GENES  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/791,495  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Van Amsterdam, John R.  
; REGISTRATION NUMBER: 40,212  
; REFERENCE/DOCKET NUMBER: L0461/7005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEFAX: 617-720-2441  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 180 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-791-495-9

Query Match 100.0%; Score 107; DB 2; Length 180;

Best Local Similarity 100.0%; Pred No. 2.1e+05; Indels 0; Gaps 0;  
Matches 20; Conservative 0; Mismatches 0;

Qy 1 RGRGAGAAASGGGAPR 20

Db 43 RGRGAGAAASGGGAPR 62

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 180  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-937-263B-8

Query Match 94.4%; Score 101; DB 3; Length 180;  
Best Local Similarity 95.0%; Pred. No. 9.9e-05;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RGRPGAGAAASGGGGAPR 20  
DB 43 RAPRGAGAAASGGGGAPR 62

RESULT 4  
US-09-751-798-8  
Sequence 8, Application US/09751798  
Patent No. 6525177  
GENERAL INFORMATION:  
APPLICANT: Stockert, Elisabeth; Jager, Elke;  
APPLICANT: Chen, Yao-tseng; Scanlan, Matthew;  
APPLICANT: Knuth, Alexander; Old, Lloyd J.  
TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer  
TITLE OF INVENTION: Associated Proteins, Uses Thereof,  
TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA  
TITLE OF INVENTION: Binding Peptides Derived Therefrom  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski, L.L.P.  
STREET: 666 Fifth Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: WordPerfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/751,798  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/062,422  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/725,182  
FILING DATE: October 3, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6525177man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5466.3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 318-3169  
TELEFAX: (212) 752-5958  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 180  
TYPE: amino acid  
TOPOLOGY: linear  
US-09-751-798-8

Query Match 94.4%; Score 101; DB 4; Length 180;  
Best Local Similarity 95.0%; Pred. No. 9.9e-05;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RGRPGAGAAASGGGGAPR 20  
DB 43 RAPRGAGAAASGGGGAPR 62

US-09-392-714-25  
Sequence 25, Application US/09392714A  
Patent No. 6686147  
GENERAL INFORMATION:  
APPLICANT: Scanlan, Matthew J.  
APPLICANT: Gure, Ali O.  
APPLICANT: Williamson, Barbara  
APPLICANT: Chen, Yao-Tseng  
APPLICANT: Old, Lloyd J.  
TITLE OF INVENTION: Cancer Associated Antigens and Uses  
TITLE OF INVENTION: Therefor  
FILE REFERENCE: L0461/7062  
CURRENT APPLICATION NUMBER: US/09/392.714A  
EARLIER FILING DATE: 1999-09-09  
EARLIER APPLICATION NUMBER: PCT/US98/14679  
EARLIER FILING DATE: 1998-07-15  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: Fast-Seq for Windows Version 3.0  
SEQ ID NO 25  
LENGTH: 180  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-392-714-25

Query Match 100.0%; Score 107; DB 4; Length 180;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGRPGAGAAASGGGGAPR 20  
DB 43 RGRPGAGAAASGGGGAPR 62

RESULT 3  
US-08-937-263B-8  
Sequence 8, Application US/08937263B  
Patent No. 6274145  
GENERAL INFORMATION:  
APPLICANT: Chen, Yao-Tseng; Scanlan, Matthew;  
APPLICANT: Gure, Ali; Old, Lloyd J.; Jager, Elke;  
APPLICANT: Alexander, Knuth; Drijfhout, Jan W.  
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE  
TITLE OF INVENTION: ENCODING CANCER ASSOCIATED ANTIGEN, THE ANTIGEN  
TITLE OF INVENTION: ITSELF, AND USES THEREOF  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski, L.L.P.  
STREET: 666 Fifth Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: WordPerfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/937.263B  
FILING DATE: September 15, 1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/725,182  
FILING DATE: October 3, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Sidd, Eric, Patent Agent  
REGISTRATION NUMBER: 40,177  
REFERENCE/DOCKET NUMBER: LUD 5466.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 318-3000  
TELEFAX: (212) 752-5958

```
RESULT 5
US-08-791-495-7
; Sequence 7, Application US/08791495
; Patent No. 5811519
; GENERAL INFORMATION:
; APPLICANT: Leth, Bernard
; APPLICANT: Lucas, Sophie
; APPLICANT: De Smet, Charles
; APPLICANT: Godelaine, Daniele
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/791,495
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Van Amsterdam, John R.
; REGISTRATION NUMBER: 40,212
; REFERENCE/DOCKET NUMBER: L0461/7005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-791-495-7

Query Match 92.5%; Score 99; DB 2; Length 180;
Best Local Similarity 95.0%; Pred. No. 0.00017;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RGPRGAGARASGPGGAPR 20
Db 43 RGPRGAGARASGPRGGAPR 62

RESULT 6
US-08-791-495-5
; Sequence 5, Application US/08791495
; Patent No. 5811519
; GENERAL INFORMATION:
; APPLICANT: Leth, Bernard
; APPLICANT: Lucas, Sophie
; APPLICANT: De Smet, Charles
; APPLICANT: Godelaine, Daniele
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/791,495
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Van Amsterdam, John R.
; REGISTRATION NUMBER: 40,212
; REFERENCE/DOCKET NUMBER: L0461/7005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-791-495-5

Query Match 92.5%; Score 99; DB 2; Length 180;
Best Local Similarity 95.0%; Pred. No. 0.00017;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RGPRGAGARASGPGGAPR 20
Db 43 RGPRGAGARASGPRGGAPR 62

RESULT 7
US-09-252-991A-25681
; Sequence 25681, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25681
; LENGTH: 860
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25681

Query Match 58.9%; Score 63; DB 4; Length 860;
Best Local Similarity 63.2%; Pred. No. 8.2;
Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RGPRGAGARASGPGGAP 19
Db 738 RKPRGSTAARQTGGGGTTP 756

RESULT 8
US-09-252-991A-27091
; Sequence 27091, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
```

US-08-440-815-2  
; Sequence 2, Application US/08440815  
; Patent No. 5798448  
; GENERAL INFORMATION:  
; APPLICANT: Caras, Ingrid W.  
; APPLICANT: Winslow, John W.  
; TITLE OF INVENTION: AL-1 Neurotrophic Factor  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/440,815  
; FILING DATE: 15-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/330128  
; FILING DATE: 27-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Torchia, Timothy E.  
; REGISTRATION NUMBER: 36,700  
; REFERENCE/DOCKET NUMBER: 920C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-8674  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 928 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-440-815-2  
Query Match 56.1%; Score 60; DB 1; Length 928;  
Best Local Similarity 61.9%; Pred. No. 19;  
Matches 13; Conservative 0; Mismatches 6; Indels 2; Gaps 1;  
QY 2 GPRGAGAAASGPGGG--APR 20  
|||||  
Db 5 GPRGAGRRRTGRRGGGDTPT 25  
RESULT 11  
US-08-486-449-2  
; Sequence 2, Application US/08486449  
; Patent No. 6280732  
; GENERAL INFORMATION:  
; APPLICANT: Caras, Ingrid W.  
; APPLICANT: Winslow, John W.  
; TITLE OF INVENTION: AL-1 Neurotrophic Factor  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:

PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 27091  
; LENGTH: 160  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27091  
Query Match 56.1%; Score 60; DB 4; Length 160;  
Best Local Similarity 63.6%; Pred. No. 3.8;  
Matches 14; Conservative 0; Mismatches 4; Indels 4; Gaps 1;  
QY 3 PRGAGARASG----PGGGAPR 20  
|||||  
Db 103 PRGAGARASGRCRRRPPR 124  
RESULT 9  
US-08-442-248-2  
; Sequence 2, Application US/08442248  
; Patent No. 5759663  
; GENERAL INFORMATION:  
; APPLICANT: Caras, Ingrid W.  
; APPLICANT: Winslow, John W.  
; TITLE OF INVENTION: AL-1 Neurotrophic Factor  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/442,248  
; FILING DATE: 15-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/330128  
; FILING DATE: 27-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Torchia, Timothy E.  
; REGISTRATION NUMBER: 36,700  
; REFERENCE/DOCKET NUMBER: 920C4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-8674  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 928 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-442-248-2  
Query Match 56.1%; Score 60; DB 1; Length 928;  
Best Local Similarity 61.9%; Pred. No. 19;  
Matches 13; Conservative 0; Mismatches 6; Indels 2; Gaps 1;  
QY 2 GPRGAGAAASGPGGG--APR 20  
|||||  
Db 5 GPRGAGRRRTGRRGGGDTPT 25  
RESULT 10

```

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/578,684
FILING DATE: 02-Jan-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/330128
FILING DATE: 27-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/486449
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Phd., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P0920P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 928 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear

```

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/578,684  
FILING DATE: 02-Jan-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/330128  
FILING DATE: 27-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/486449  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, Phd., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P0920P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 928 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear

```

REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1005 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-469-537A-103

Query Match          56.1%; Score 60; DB 2;
Best Local Similarity 61.9%; Pred. No. 21;
Matches 13; Conservative

QY      2  GPRGAGAAASGPGGG--APR 20
          ||||| | | ||| ||
Ddb     5  GPRGAGRRRTQGRGGGDTPR 25

RESULT 14
US-09-252-991A-24923
; Sequence 24923, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

```

APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCES: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 24923  
LENGTH: 210  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24923

Query Match 55.1%; Score 59; DB 4; Length 210;

Best Local Similarity 70.6%; Pred. No. 6.4;

Matches 12; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 RGAGAAAGGGGGGAPR 20  
DB 154 RPAAGTAAAGGGGGAPR 170

#### RESULT 15

US-08-118-270-21

Sequence 21, Application US/08118270

Patent No. 5508384

GENERAL INFORMATION:

APPLICANT: Murphy, Randall B.

APPLICANT: Schuster, David I.

TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN

TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF

NUMBER OF SEQUENCES: 348

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK

STREET: 419 Seventh Street, N.W., Suite 300

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/118,270

FILING DATE: 09-SEP-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/943,236

FILING DATE: 10-SEP-1992

ATTORNEY/AGENT INFORMATION:

NAME: Townsend, Kevin G.

REGISTRATION NUMBER: 34,033

REFERENCE/DOCKET NUMBER: MURPHY-2A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528

TELEX: 248633

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 330 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-118-270-21

Query Match

Best Local Similarity

55.1%; Score 59; DB 1; Length 330;

55.0%; Pred. No. 9.6;

Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;  
QY 1 RGPRGAGAAAGGGGGGAPR 20  
DB 212 RRPNGLGPRGAGGGGGR 231

Search completed: June 22, 2004, 08:54:48

Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2004, 08:51:28 ; Search time 52 Seconds  
(without alignments)  
108.582 Million cell updates/sec

Title: US-09-529-206D-4\_COPY\_43\_62

Perfect score: 107  
Sequence: 1 RGPAGAGARASPGGGAPR 20

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1163542 seqs, 282313646 residues

Total number of hits satisfying chosen parameters: 1163542

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	107	100.0	180	10	US-09-849-602-30
2	107	100.0	180	12	US-10-296-734-832
3	107	100.0	180	14	US-10-207-655-71
4	107	100.0	180	15	US-10-026-066-3
5	107	100.0	180	15	US-10-117-937-74
6	107	100.0	180	15	US-10-295-027-386
7	107	100.0	180	16	US-10-188-832-139
8	107	100.0	397	9	US-09-821-883-27
9	102	95.3	30	12	US-10-296-734-1404
10	102	95.3	3541	12	US-10-296-734-1454
11	101	94.4	180	9	US-09-751-798-8
12	101	94.4	180	13	US-10-023-182-8
13	101	94.4	180	14	US-10-364-614-14
14	99	92.5	135	15	US-10-295-027-388
15	99	92.5	135	16	US-10-188-832-141

Sequence 834, Appl  
Sequence 69, Appl  
Sequence 75, Appl  
Sequence 88, Appl  
Sequence 76, Appl  
Sequence 496, Appl  
Sequence 1436, Appl  
Sequence 1402, Appl  
Sequence 1434, Appl  
Sequence 26, Appl  
Sequence 27, Appl  
Sequence 2, Appl  
Sequence 63, Appl  
Sequence 34, Appl  
Sequence 468, Appl  
Sequence 468, Appl  
Sequence 53847, A  
Sequence 18, Appl  
Sequence 48480, A  
Sequence 43408, A  
Sequence 1480, Appl  
Sequence 1188, Appl  
Sequence 142, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 4, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 65, Appl  
Sequence 2770, Appl

16 99 92.5 180 12 US-10-296-734-834  
17 99 92.5 180 14 US-10-146-473-69  
18 99 92.5 180 15 US-10-117-937-75  
19 99 92.5 210 14 US-10-157-031-88  
20 99 92.5 210 15 US-10-117-937-76  
21 95 88.8 20 15 US-10-313-986-496  
22 94 87.9 30 12 US-10-296-734-1436  
23 85 79.4 30 12 US-10-296-734-1402  
24 77 72.0 30 12 US-10-296-734-1434  
25 60 56.1 450 14 US-10-001-073-26  
26 60 56.1 450 14 US-10-001-073-27  
27 60 56.1 928 8 US-08-578-684-2  
28 60 56.1 1005 12 US-10-029-020-63  
29 59 55.1 123 16 US-10-408-765A-34  
30 59 55.1 159 9 US-09-925-302-468  
31 59 55.1 159 12 US-09-925-302-468  
32 58 54.2 186 12 US-10-425-114-53847  
33 57 53.3 254 12 US-10-343-348-18  
34 57 53.3 278 12 US-10-425-114-48480  
35 57 53.3 279 12 US-10-425-114-43608  
36 57 53.3 280 14 US-10-017-161-1480  
37 57 53.3 280 15 US-10-292-798-1188  
38 57 53.3 418 12 US-10-262-839-142  
39 57 53.3 865 10 US-09-842-758-20  
40 57 53.3 865 12 US-10-174-333-20  
41 57 53.3 866 12 US-10-433-579-4  
42 57 53.3 934 10 US-09-842-758-18  
43 57 53.3 934 12 US-10-174-333-18  
44 56 52.3 256 14 US-10-301-822-65  
45 56 52.3 256 15 US-10-094-749-2770

## ALIGNMENTS

RESULT 1  
US-09-849-602-30  
; Sequence 30, Application US/09849602  
; Publication No. US20030165834A1  
; GENERAL INFORMATION:  
; APPLICANT: Scanlan, Matthew J.  
; APPLICANT: Old, Lloyd J.  
; APPLICANT: Stockert, Elisabeth  
; APPLICANT: Chen, Yao-Tseung  
; TITLE OF INVENTION: Colon Cancer Antigen Panel  
; FILE REFERENCE: L0461/7105(JRV)  
; CURRENT APPLICATION NUMBER: US/09/849,602  
; CURRENT FILING DATE: 2001-05-04  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 30  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-849-602-30

Query Match 100.0%; Score 107; DB 10; Length 180;  
Best Local Similarity 100.0%; Pred. No. 0.00033;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGPAGAGARASPGGGAPR 20  
|||||  
DB 43 RGPAGAGARASPGGGAPR 62

RESULT 2  
US-10-296-734-832  
; Sequence 832, Application US/10296734  
; Publication No. US20040054137A1  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Scott A  
; APPLICANT: Ramshaw, Ian A  
; TITLE OF INVENTION: Synthetic molecules and uses therefor

```
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU P07761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 832
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: NYNsola consensus polypeptide
US-10-296-734-832

Query Match      100.0%; Score 107; DB 12; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGRGAGAAARASGGGGAPR 20
Db      43 RGRGAGAAARASGGGGAPR 62

RESULT 3
US-10-207-655-71
; Sequence 71, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 71
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-71

Query Match      100.0%; Score 107; DB 14; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGRGAGAAARASGGGGAPR 20
Db      43 RGRGAGAAARASGGGGAPR 62

RESULT 4
US-10-026-066-3
; Sequence 3, Application US/10026066
; Publication No. US20030215425A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; TITLE OF INVENTION: EPITOPE SYNCHRONIZATION IN ANTIGEN
; PRESENTING CELLS
; FILE REFERENCE: CTLLM.21C1C
; CURRENT APPLICATION NUMBER: US/10/026,066
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 09/561,074
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/560,465
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/561,572
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/561,571
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: PCT/US01/13806

; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU P07761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 832
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: NYNsola consensus polypeptide
US-10-296-734-832

Query Match      100.0%; Score 107; DB 15; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGRGAGAAARASGGGGAPR 20
Db      43 RGRGAGAAARASGGGGAPR 62

RESULT 5
US-10-117-937-74
; Sequence 74, Application US/10117937
; Publication No. US20030220239A1
; GENERAL INFORMATION:
; APPLICANT: CTL IMMUNO THERAPIES CORP.
; APPLICANT: SIMARD, John, J.L.
; APPLICANT: DIAMOND, David, C.
; APPLICANT: LIU, Liping
; APPLICANT: XIE, Zhidong
; TITLE OF INVENTION: EPITOPE SEQUENCES
; FILE REFERENCE: CTLLM.027A
; CURRENT APPLICATION NUMBER: US/10/117,937
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/282,211
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/337,017
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/363,210
; NUMBER OF SEQ ID NOS: 602
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-937-74

Query Match      100.0%; Score 107; DB 15; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGRGAGAAARASGGGGAPR 20
Db      43 RGRGAGAAARASGGGGAPR 62

RESULT 6
US-10-295-027-386
; Sequence 366, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
```

```
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US/09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US/60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US/60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US/60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US/60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US/60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US/60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US/60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US/60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US/60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 386
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-386

Query Match      100.0%; Score 107; DB 15; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPRGAGARASGPGGAPR 20
Db 43 RPRGAGARASGPGGAPR 62

RESULT 7
US-10-188-832-139
; Sequence 139, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US/60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US/60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US/60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US/60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US/60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 139
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-832-139

Query Match      100.0%; Score 107; DB 16; Length 180;

Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPRGAGARASGPGGAPR 20
Db 43 RPRGAGARASGPGGAPR 62

US-09-821-883-27
; Sequence 27, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Thomas
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; TITLE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US/60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: NY-R50-IC tumor antigen
US-09-821-883-27

Query Match      100.0%; Score 107; DB 9; Length 397;
Best Local Similarity 100.0%; Pred. No. 0.00055;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPRGAGARASGPGGAPR 20
Db 43 RPRGAGARASGPGGAPR 62

RESULT 9
US-10-296-734-1404
; Sequence 1404, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: Patent In version 3.2
; SEQ ID NO 1404
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: NYS01a segment 4
US-10-296-734-1404

Query Match      95.3%; Score 102; DB 12; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GPRGAGARASGPGGAPR 20
Db 1 GPRGAGARASGPGGAPR 19
```

RESULT 10  
US-10-296-734-1454  
; Sequence 1454, Application US/10296734  
; Publication No. US20040054137A1  
; GENERAL INFORMATION:  
; APPLICANT: Ramshaw, Ian A  
; APPLICANT: Thompson, Scott A  
; TITLE OF INVENTION: Synthetic molecules and uses therefor  
; FILE REFERENCE: Savine  
; CURRENT APPLICATION NUMBER: US/10/296.734  
; CURRENT FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: AU P07761/00  
; PRIOR FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 1507  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1454  
; LENGTH: 3541  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Melanoma cancer specific savine  
US-10-296-734-1454

Query Match 95.3%; Score 102; DB 12; Length 3541;  
Best Local Similarity 100.0%; Pred. No. 0.015;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GPRGAGARASGPGGAPR 20  
Db 2359 GPRGAGARASGPGGAPR 2377

RESULT 11  
US-09-751-798-8  
; Sequence 8, Application US/09751798  
; Patent No. US20020010321A1  
; GENERAL INFORMATION:  
; APPLICANT: Stockert, Elisabeth; Jager, Elke;  
; APPLICANT: Chen, Yao-tseng; Scanlan, Matthew;  
; APPLICANT: Knuth, Alexander; Old, Lloyd J.  
; TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer  
; TITLE OF INVENTION: Associated Proteins, Uses Thereof,  
; TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA  
; TITLE OF INVENTION: Binding Peptides Derived Therefrom  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Fulbright & Jaworski, L.L.P.  
; STREET: 666 Fifth Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/751,798  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/062,422  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/725,182  
; FILING DATE: October 3, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. US20020010321A1man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: LUD 5466.3  
; TELEPHONE: (212) 318-3168  
; TELECOMMUNICATION INFORMATION:

TELEFAX: (212) 752-5958  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 180  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-09-751-798-8  
Query Match 94.4%; Score 101; DB 9; Length 180;  
Best Local Similarity 95.0%; Pred. No. 0.0014;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 RGRGAGARASGPGGAPR 20  
Db 43 RAPRGAGARASGPGGAPR 62  
RESULT 12  
US-10-023-182-8  
; Sequence 8, Application US/10023182  
; Publication No. US20020164665A1  
; GENERAL INFORMATION:  
; APPLICANT: Stockert, Elisabeth; Jager, Elke;  
; Chen, Yao-tseng; Scanlan, Matthew;  
; Knuth, Alexander; Old, Lloyd J.  
; TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer  
; TITLE OF INVENTION: Associated Proteins, Uses Thereof,  
; TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA  
; TITLE OF INVENTION: Binding Peptides Derived Therefrom  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Fulbright & Jaworski, L.L.P.  
; STREET: 666 Fifth Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/023,182  
; FILING DATE: 17-Dec-2001  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/751,798  
; FILING DATE: December 29, 2000  
; APPLICATION NUMBER: 09/062,422  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: 08/725,182  
; FILING DATE: October 3, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. US20020164665A1man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: LUD 5466.3  
; TELEPHONE: (212) 318-3168  
; TELEFAX: (212) 752-5958  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 180  
; TYPE: amino acid  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-10-023-182-8  
Query Match 94.4%; Score 101; DB 13; Length 180;  
Best Local Similarity 95.0%; Pred. No. 0.0014;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 RGRGAGARASGPGGAPR 20

Db 43 RAPRGAGAARASGPGGAPR 62

## RESULT 13

US-10-364-614-14  
; Sequence 14, Application US/10364614  
; Publication No. US20030175250A1  
; GENERAL INFORMATION:  
; APPLICANT: JAGER, Elke  
; APPLICANT: KNUTH, Alexander  
; APPLICANT: OLD, Lloyd  
; APPLICANT: Gnjatic, Sacha  
; TITLE OF INVENTION: ISOLATED PEPTIDES WHICH BIND TO HLA MOLECULES AND USES THEREOF  
; FILE REFERENCE: LUD 5726.1 CIP US/10/364,614  
; CURRENT APPLICATION NUMBER: US/10/364,614  
; CURRENT FILING DATE: 2003-02-24  
; PRIOR APPLICATION NUMBER: US 60/355,828  
; PRIOR FILING DATE: 2002-02-13  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 14  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-364-614-14

Query Match 94.4%; Score 101; DB 14; Length 180;  
Best Local Similarity 95.0%; Pred. No. 0.0014;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RGRGAGAARASGPGGAPR 20  
Db 43 RAPRGAGAARASGPGGAPR 62

## RESULT 14

US-10-295-027-388  
; Sequence 388, Application US/10295027  
; Publication No. US20030232350A1  
; GENERAL INFORMATION:  
; APPLICANT: Afar, Daniel  
; APPLICANT: Aziz, Nataasha  
; APPLICANT: Ginsberg, Wendy M.  
; APPLICANT: Gish, Kurt C.  
; APPLICANT: Glynn, Richard  
; APPLICANT: Hevezi, Peter A.  
; APPLICANT: Mack, David H.  
; APPLICANT: Murray, Richard  
; APPLICANT: Watson, Susan R.  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
; FILE REFERENCE: 018501-012500US  
; CURRENT APPLICATION NUMBER: US/10/295,027  
; CURRENT FILING DATE: 2002-11-13  
; PRIOR APPLICATION NUMBER: US 09/663,733  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/335,394  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/332,464  
; PRIOR FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US 60/334,393  
; PRIOR FILING DATE: 2001-11-29  
; PRIOR APPLICATION NUMBER: US 60/340,376  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: US 60/347,211  
; PRIOR FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 60/347,349  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/355,250  
; PRIOR FILING DATE: 2002-02-08

; PRIOR APPLICATION NUMBER: US 60/356,714  
; PRIOR FILING DATE: 2002-02-13  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1386  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 388  
; LENGTH: 135  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-295-027-388

Query Match 92.5%; Score 99; DB 15; Length 135;  
Best Local Similarity 95.0%; Pred. No. 0.0018;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RGRGAGAARASGPGGAPR 20  
Db 43 RGRGAGAARASGPGGAPR 62

## RESULT 15

US-10-188-832-141  
; Sequence 141, Application US/10188832  
; Publication No. US20040076955A1  
; GENERAL INFORMATION:  
; APPLICANT: Mack, David H.  
; APPLICANT: Aziz, Nataasha  
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions  
; FILE REFERENCE: 018501-002330US  
; CURRENT APPLICATION NUMBER: US/10/188,832  
; CURRENT FILING DATE: 2002-11-22  
; PRIOR APPLICATION NUMBER: US 60/302,814  
; PRIOR FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: US 60/310,099  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: US 60/343,705  
; PRIOR FILING DATE: 2001-11-08  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/372,246  
; PRIOR FILING DATE: 2002-04-12  
; NUMBER OF SEQ ID NOS: 207  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 141  
; LENGTH: 135  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-188-832-141

Query Match 92.5%; Score 99; DB 16; Length 135;  
Best Local Similarity 95.0%; Pred. No. 0.0018;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RGRGAGAARASGPGGAPR 20  
Db 43 RGRGAGAARASGPGGAPR 62

Search completed: June 22, 2004, 08:56:18  
Job time : 52 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2004, 08:48:03 ; Search time 18 Seconds  
(without alignments)  
106.879 Million cell updates/sec

Title: US-09-529-206d-4\_COPY\_43\_62  
Perfect score: 107  
Sequence: 1 RGRGAGARASGPGGAPR 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

PIR 78: \*  
1: PIR1: \*  
2: PIR2: \*  
3: PIR3: \*  
4: PIR4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	62.6	627	2 T35608	polyketide hydroxy
2	60	56.1	893	2 S51603	receptor-like tyro
3	60	56.1	981	2 S51604	receptor-like tyro
4	60	56.1	1005	2 S49015	receptor tyrosine
5	59	55.1	129	1 OTHU5H	cytochrome-c oxida
6	59	55.1	877	2 I48967	brain-specific kin
7	58.5	54.7	566	2 T21096	hypothetical prote
8	58.5	54.7	692	2 T21095	hypothetical prote
9	57	53.3	129	1 JC2254	cytochrome-c oxida
10	57	53.3	355	2 T36273	hypothetical prote
11	56	52.3	331	2 S78452	POU-domain protein
12	55	51.4	201	2 C40040	alternative splici
13	55	51.4	201	2 S26404	alternative splici
14	55	51.4	248	2 A40040	alternative splici
15	55	51.4	250	2 T46417	hypothetical prote
16	55	51.4	292	2 B40040	alternative splici
17	55	51.4	1215	2 T32734	myosin-1A - Acanth
18	54.5	50.9	496	2 F83124	probable transcrip
19	54	50.5	388	2 T29173	hypothetical prote
20	54	50.5	537	2 S51212	BAK5 protein - bov
21	54	50.5	1730	2 A36226	collagen alpha 1 c
22	54	50.5	1733	1 B45344	probable nuclear a
23	53.5	50.0	627	2 T41112	spidroin 2, dragli
24	53.5	50.0	3150	2 T13828	CREB-binding prote
25	53	49.5	485	1 F71275	probable periplasm
26	53	49.5	1306	2 A70934	hypothetical glyci
27	52.5	49.1	1450	2 A34169	alpha-2A-adrenogl
28	52.5	49.1	782	2 S33945	late protein, 100K
29	52.5	49.1	1690	2 T35694	ATP dependent DNA

RESULT 1

T35608  
polyketide hydroxylase - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C>Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 17-Mar-2000  
C:Accession: T35608  
R:Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, June 1999  
A:Reference number: Z21584  
A:Accession: T35608  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-627 <SEE>  
A:Cross-references: EMBL:AL079356; PIDN:CAB45603.1; GSPDB:GN00070; SCODEB:SC6G9.12c  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCODEB:SC6G9.12c  
C:Superfamily: tetracycline 6-hydroxylase

Query Match 62.6%; Score 67; DB 2; Length 627;  
Best Local Similarity 75.0%; Pred. No. 2;  
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GPRGAGARASGPGG 17  
DB 447 GPRGAGGAPGGPGG 462

RESULT 2

S51603  
receptor-like tyrosine kinase Etk-1 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 28-May-1999  
C:Accession: S51603  
R:Maizsonpierre, P.C.; Barrezaeta, N.X.; Yancopoulos, G.D.  
Oncogene 8, 3277-3288, 1993  
A>Title: Etk-1 and Etk-2; two novel members of the Eph receptor-like tyrosine kinase fa  
A:Reference number: S49015; MUID:94067777; PMID:7504232  
A:Accession: S51603  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-893 <MAI>  
A:Cross-references: EMBL:S68028  
A>Note: the authors translated the codon GAC for residue 170 as Glu  
C:Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat  
C:Keywords: ATP; transmembrane protein  
F:563-829/Domain: protein kinase homology <KIN>  
F:571-579/Region: protein kinase ATP-binding motif

Query Match 56.1%; Score 60; DB 2; Length 893;  
Best Local Similarity 61.9%; Pred. No. 16;  
Matches 13; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

ALIGNMENTS

30 52.5 49.1 1737 2 T00209  
31 52 48.6 295 2 T04483  
32 52 48.6 314 2 JC5273  
33 52 48.6 373 2 A47234  
34 52 48.6 676 1 EDBE22  
35 52 48.6 676 1 EDBE23  
36 52 48.6 1171 2 T42372  
37 52 48.6 1256 2 JE0209  
38 52 48.6 1487 1 EDBE21  
39 52 48.6 1487 1 EDBE22  
40 51.5 48.1 1958 2 B40505  
41 51 47.7 198 2 A57717  
42 51 47.7 327 2 T47231  
43 51 47.7 420 2 I59234  
44 51 47.7 632 2 S42731  
45 51 47.7 638 2 T48796

MEGF8 protein - hu  
probable ring fing  
paired type homeob  
homeobox protein H  
immediate-early pr  
immediate-early pr  
probable guanylate  
brain-specific ang  
immediate-early pr  
155K transcription  
hypothetical prote  
transcription fact  
fibrillarlin (impor  
octamer binding tr  
collagen alpha 1 c  
probable ATP-depen

QY 2 GPRGAGARASGPGGG--APR 20  
 ||||| | | | | |  
 Db 5 GPRGAGRRRTQRCGGGGTTPR 25

RESULT 3  
 S51604  
 A:Title: receptor-like tyrosine kinase Etk-1 - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 29-May-1998  
 C:Accession: S51604  
 R:Maizonnier, P.C.; Barreuzeta, N.X.; Yancopoulos, G.D.  
 Oncogene 8, 3277-3288, 1993  
 A:Title: Etk-1 and Etk-2: two novel members of the Eph receptor-like tyrosine kinase fam  
 A:Reference number: S49015; MUID:94067777; PMID:7504232  
 A:Accession: S51604  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-381 <MAI>  
 A:Cross-references: EMBL:S68029  
 A:Note: the authors translated the codon GAC for residue 170 as Glu  
 C:Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h  
 C:Keywords: ATP; transmembrane protein  
 F:651-917/Domain: protein kinase homology <KIN>  
 F:659-667/Region: protein kinase ATP-binding motif

Query Match 56.1%; Score 60; DB 2; Length 981;  
 Best Local Similarity 61.9%; Pred. No. 17;  
 Matches 13; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY 2 GPRGAGARASGPGGG--APR 20  
 ||||| | | | | |  
 Db 5 GPRGAGRRRTQRCGGGGTTPR 25

RESULT 4  
 S49015  
 A:Title: receptor tyrosine kinase Etk-1 - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 14-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 28-May-1999  
 C:Accession: S49015; S51602  
 R:Maizonnier, P.C.; Barreuzeta, N.X.; Yancopoulos, G.D.  
 Oncogene 8, 3277-3288, 1993  
 A:Title: Etk-1 and Etk-2: two novel members of the Eph receptor-like tyrosine kinase fam  
 A:Reference number: S49015; MUID:94067777; PMID:7504232  
 A:Accession: S49015  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1005 <MAI>  
 A:Cross-references: EMBL:S68024  
 A:Note: the authors translated the codon GAC for residue 170 as Glu  
 A:Accession: S51602  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-305, 'G', 359-1005 <MA2>  
 A:Cross-references: EMBL:S68026  
 A:Note: the authors translated the codon GAC for residue 170 as Glu  
 C:Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h  
 C:Keywords: ATP; transmembrane protein  
 F:675-941/Domain: protein kinase homology <KIN>  
 F:683-691/Region: protein kinase ATP-binding motif

Query Match 56.1%; Score 60; DB 2; Length 1005;  
 Best Local Similarity 61.9%; Pred. No. 18;  
 Matches 13; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY 2 GPRGAGARASGPGGG--APR 20  
 ||||| | | | | |  
 Db 5 GPRGAGRRRTQRCGGGGTTPR 25

RESULT 5

OTHU58  
 cytochrome-c oxidase (EC 1.9.3.1) chain Vb precursor - human  
 C:Species: Homo sapiens (man)  
 C>Date: 31-Mar-1992 #sequence\_revision 07-Jun-1996 #text\_change 11-Jun-1999  
 C:Accession: J0324; A39063; S74198; A28817  
 R:Zeviani, M.; Sakoda, S.; Sherbany, A.A.; Nakase, H.; Rizzuto, R.; Samitt, C.E.; DiMat  
 Gene 65, 1-11, 1988  
 A:Title: Sequence of cDNAs encoding subunit Vb of human and bovine cytochrome c oxidase  
 A:Reference number: J0324; MUID:98284368; PMID:2840351  
 A:Accession: J0324  
 A:Molecule type: mRNA  
 A:Residues: 1-129 <ZEV>  
 A:Cross-references: EMBL:M19961; NID:G180940; PIDN:AAA52061.1; PID:G180941  
 R:Lomax, M.I.; Hsieh, C.L.; Darvas, B.T.; Francke, U.  
 Genomics 10, 1-9, 1991  
 A:Title: Structure of the human cytochrome c oxidase subunit Vb gene and chromosomal ma  
 A:Reference number: A39063; MUID:91257815; PMID:1646156  
 A:Accession: A39063  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-108, 'E', 110-129 <LOW>  
 A:Cross-references: EMBL:M59250  
 A:Note: the authors translated the codon GGC for residue 21 as His  
 R:Bachman, N.J.; Yang, T.L.; Dassen, J.S.; Ernst, R.E.; Lomax, M.I.  
 Arch. Biochem. Biophys. 333, 152-162, 1996  
 A:Title: Phylogenetic footprinting of the human cytochrome c oxidase subunit Vb promote  
 A:Reference number: S74198; MUID:96400390; PMID:8806766  
 A:Accession: S74198  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-35, 'TR' <BAC>  
 A:Cross-references: EMBL:U41284; NID:G1679627; PIDN:AAB19185.1; PID:G1679628  
 C:Genetics:  
 A:Gene: GDB:COX5B  
 A:Cross-references: GDB:127530; OMIM:123866  
 A:Map position: 2cen-2q13  
 A:Genome: nuclear  
 C:Complex: part of a 13 chain complex spanning the inner mitochondrial membrane and cor  
 (see PIR:OTHUSA), Vb, Via (see PIR:OGHUGA), Vb (see PIR:OGHUG5), Vlc (see PIR:OGHUGC)  
 m dimers within the mitochondrial inner-membrane  
 C:Function:  
 A:Description: the cytochrome-c oxidase complex catalyzes the oxidation of four molecu  
 ns from the mitochondrial matrix producing two molecules of water and lowering the conc  
 A:Note: the role of chain Vb is not clear  
 C:Superfamily: mammalian cytochrome-c oxidase chain Vb  
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrar  
 F:1-31/Domain: transit peptide (mitochondrion) #status predicted <TNP>  
 F:32-129/Product: cytochrome-c oxidase chain Vb #status predicted <MAT>  
 F:91,93,113,116/Binding site: zinc (Cys) #status predicted

Query Match 55.1%; Score 59; DB 1; Length 129;  
 Best Local Similarity 57.3%; Pred. No. 4.5;  
 Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 RGPGRGAGARASGPGGGAP 19  
 ||||| | | | | |  
 Db 20 RGPGRGAGARASGPGGGVP 38

RESULT 6  
 I48967  
 A:Title: brain-specific kinase - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 18-Jun-1999  
 C:Accession: I48967  
 R:Zhong, R.K.; Copeland, T.D.; Kroner, L.F.; Schulz, N.T.  
 J. Neurosci. Res. 37, 129-143, 1994  
 A:Title: Isolation and characterization of Bsk, a growth factor receptor-like tyrosine  
 A:Reference number: I48967; MUID:94194581; PMID:8145300  
 A:Accession: I48967  
 A:Status: preliminary; translated from GB/EMBL/DDAJ  
 A:Molecule type: mRNA  
 A:Residues: 1-877 <RES>

A;Cross-references: EMBL:U07357; NID:G466369; PIDN:AAA17038.1; PID:G466370  
C;Gene: Bax  
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat  
F:512-778/Domain: protein kinase homology <KIN>  
F:520-528/Region: protein kinase ATP-binding motif  
F:501-568/Domain: SAM homology <SAM>

Query Match 55.1%; Score 59; DB 2; Length 877;  
Best Local Similarity 61.9%; Pred. No. 20;  
Matches 13; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY 2 GPRGAGARASGPGG--GGAP 20  
DB 5 GPRGAGHRRTRGCGGDDTPR 25  
|||||

RESULT 7  
T21096  
Hypothetical protein F18H3.3b - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C;Accession: T21096  
R;Coles, L.  
Submitted to the EMBL Data Library, July 1995  
A;Reference number: Z19373  
A;Accession: T21096  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-566 <WIL>  
A;Cross-references: EMBL:Z50110; PIDN:CAA90446.1; GSPDB:GN00028; CESP:F18H3.3b  
A;Experimental source: clone F18H3  
C;Genetics:  
A;Gene: CESP:F18H3.3b  
A;Map position: X  
A;Introns: 111/1; 215/2; 469/3; 552/3  
C;Superfamily: polyadenylate-binding protein; ribonucleoprotein repeat homology

Query Match 54.7%; Score 58.5; DB 2; Length 566;  
Best Local Similarity 65.0%; Pred. No. 16;  
Matches 13; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 3 PRGAGARASGPG--GGAP 19  
DB 530 PRPAGAPRVGGPGVQMGAP 549  
|||||

RESULT 8  
T21095  
Hypothetical protein F18H3.3a - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C;Accession: T21095  
R;Coles, L.  
Submitted to the EMBL Data Library, July 1995  
A;Reference number: Z19373  
A;Accession: T21095  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-692 <WIL>  
A;Cross-references: EMBL:Z50110; PIDN:CAA90444.1; GSPDB:GN00028; CESP:F18H3.3a  
A;Experimental source: clone F18H3  
C;Genetics:  
A;Gene: CESP:F18H3.3a  
A;Map position: X  
A;Introns: 111/1; 215/2; 469/3; 552/3  
C;Superfamily: polyadenylate-binding protein; ribonucleoprotein repeat homology

Query Match 54.7%; Score 58.5; DB 2; Length 692;  
Best Local Similarity 65.0%; Pred. No. 19;  
Matches 13; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 3 PRGAGARASGPG--GGAP 19  
DB 530 PRPAGAPRVGGPGVQMGAP 549  
|||||

RESULT 9  
JC2254  
Cytochrome-c oxidase (EC 1.9.3.1) chain Vb precursor - rat  
N;Alternate names: cytochrome-c oxidase (EC 1.9.3.1) chain V1a  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 28-Oct-1994 #sequence\_revision 26-May-1995 #text\_change 16-Jun-2000  
C;Accession: JC2254; JC2255; S05318; S65375  
R;Hoshinaga, H.; Amuro, N.; Goto, Y.; Okazaki, T.  
J. Biochem. 115, 194-201, 1994  
A;Title: Molecular cloning and characterization of the rat cytochrome c oxidase subunit  
A;Reference number: JC2254; MUID:94266742; PMID:8208867  
A;Accession: JC2254  
A;Molecule type: mRNA  
A;Residues: 1-129 <HOG>  
A;Cross-references: DDBJ:D10952; NID:G493694; PIDN:BAA01744.1; PID:G493695  
A;Experimental source: liver  
A;Accession: JC2255  
A;Molecule type: DNA  
A;Residues: 1-129 <HOG>  
A;Cross-references: DDBJ:D10951  
A;Experimental source: lambda COXVb741  
R;Goto, Y.; Amuro, N.; Okazaki, T.  
Nucleic Acids Res. 17, 6388, 1989  
A;Title: Nucleotide sequence of cDNA for rat liver and brain cytochrome c oxidase subunit  
A;Reference number: S05318; MUID:89366668; PMID:2549512  
A;Accession: S05318  
A;Molecule type: mRNA  
A;Residues: 31-129 <GOT>  
A;Cross-references: EMBL:X14208  
R;Schaeffer, H.; Noack, H.; Halangk, W.; Brandt, U.; von Jagow, G.  
Eur. J. Biochem. 230, 235-241, 1995  
A;Title: Cytochrome-c oxidase in developing rat heart. Enzymic properties and amino-ten  
A;Reference number: S65372; MUID:95324529; PMID:7601105  
A;Accession: S65375  
A;Molecule type: protein  
A;Residues: 32-41 <SCH>  
A;Experimental source: liver  
C;Genetics:  
A;Gene: COXVb-1  
A;Introns: 35/1; 59/3; 93/1  
A;Note: intronless gene COXVb-2 apparently a nonfunctional processed pseudogene  
C;Superfamily: mammalian cytochrome-c oxidase chain Vb  
C;Keywords: membrane-associated complex; mitochondrial inner membrane; mitochondrion; o:  
P:1-31/Domain: transit peptide (mitochondrion) #status predicted <TNP>  
P:32-129/Product: cytochrome-c oxidase chain Vb #status predicted <MAT>

Query Match 53.3%; Score 57; DB 1; Length 129;  
Best Local Similarity 55.6%; Pred. No. 7.4;  
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 GPRGAGARASGPGGAP 19  
DB 21 GPRGAVATRSMASGGVP 38  
|||||

RESULT 10  
T36273  
Hypothetical protein SCE68.23c - Streptomyces coelicolor  
C;Species: Streptomyces coelicolor  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jun-2000  
C;Accession: T36273  
R;Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
Submitted to the EMBL Data Library, June 1999  
A;Reference number: Z21576  
A;Accession: T36273  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-355 <MUR>

A:Cross-references: EMBL:AL079345; PIDN:CAB45359.1; GSPDB:GN00070; SCOEDB:SC68.23c  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SC68.23c  
C:Superfamily: Streptomyces coelicolor hypothetical protein SC68.23c

Query Match 53.3%; Score 57; DB 2; Length 355;  
Best Local Similarity 57.9%; Pred. No. 17;  
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 RGRGAGAAARASGPGGGAP 19  
DB 15 RGRGAGAAARASGPGGGAP 33

RESULT 11  
S78452  
POU-domain protein rdc-1 - human  
C:Species: Homo sapiens (man)  
C:Date: 19-Mar-1998 #sequence\_revision 19-Mar-1998 #text\_change 24-Sep-1999  
C:Accession: S78452; S26063; I38151  
R:Alt, F.W.  
Submitted to the EMBL Data Library, February 1992  
A:Reference number: S78452  
A:Accession: S78452  
A:Molecule type: mRNA  
A:Residues: 1-331 <ALT>  
A:Cross-references: EMBL:X64624; NID:q35914; PIDN:CAA45907.1; PID:q35915  
R:Collum, R.G.; Fisher, P.E.; Datta, M.; Mellis, S.; Thiele, C.; Huebner, K.; Croce, C.M.  
Nucleic Acids Res. 26, 4919-4925, 1992  
A:Title: A novel POU homeodomain gene specifically expressed in cells of the developing  
A:Reference number: I38151; MUID:93027214; PMID:1357630  
A:Accession: S26063  
A:Molecule type: mRNA  
A:Residues: 174-177, 'S', 179-233, 'K', 235-327 <COL>  
A:Cross-references: EMBL:X64624  
A:Experimental source: placenta  
C:Superfamily: unassigned homeobox proteins; homeobox homology; POU domain homology  
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F:181-250/Domain: POU domain homology <POU>  
F:268-324/Domain: homeobox homology <HOX>

Query Match 52.3%; Score 56; DB 2; Length 331;  
Best Local Similarity 61.1%; Pred. No. 20;  
Matches 11; Conservative 0; Mismatches 3; Indels 4; Gaps 1;

QY 2 GPRGAGAAARASGPGGGAP 19  
DB 65 GPRGGG---GGPGGGP 78

RESULT 12  
C40040  
alternative splicing factor ASF-3 - human  
C:Species: Homo sapiens (man)  
C:Date: 24-Jan-1992 #sequence\_revision 24-Jan-1992 #text\_change 03-Dec-1999  
C:Accession: C40040  
R:Ge, H.; Zuo, P.; Manley, J.L.  
Cell 66, 373-382, 1991  
A:Title: Primary structure of the human splicing factor ASF reveals similarities with D  
A:Reference number: A40040; MUID:91309149; PMID:1855257  
A:Accession: C40040  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-201 <GEA>  
A:Cross-references: GB:M72709  
C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprote  
F:17-81/Domain: ribonucleoprotein repeat homology <RRM1>  
F:122-182/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 51.4%; Score 55; DB 2; Length 201;  
Best Local Similarity 55.0%; Pred. No. 18;  
Matches 11; Conservative 0; Mismatches 9; Indels 9; Gaps 0;

QY 1 RGRGAGAAARASGPGGGAP 20  
DB 90 RSGRGTGRGGGGGGGAPR 109

## RESULT 13

S26404

alternative splicing factor ASF - mouse

C:Species: Mus musculus (house mouse)

C:Date: 25-Feb-1994 #sequence\_revision 26-May-1995 #text\_change 03-Dec-1999

C:Accession: S26404

R:Tacke, R.; Boned, A.; Goridis, C.

Nucleic Acids Res. 20, 5482, 1992

A:Title: ASF alternative transcripts are highly conserved between mouse and man.

A:Reference number: S26404; MUID:93065226; PMID:1437571

A:Accession: S26404

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-201 &lt;TAC&gt;

A:Cross-references: EMBL:X66091

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1992

C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprote

F:17-81/Domain: ribonucleoprotein repeat homology &lt;RRM1&gt;

F:122-182/Domain: ribonucleoprotein repeat homology &lt;RRM2&gt;

Search completed: June 22, 2004, 08:54:02  
Job time : 19 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OK protein - protein search, using sw model

Run on: June 22, 2004, 08:41:07 ; Search time 14 Seconds  
(without alignments)  
74.386 Million cell updates/sec

Title: US-09-529-206D-4\_COPY\_43\_62  
Perfect score: 107  
Sequence: 1 RGPAGAGARASGGGGR 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	180	1	CTG1_HUMAN
2	99	92.5	210	1	CTG2_HUMAN
3	67	62.6	627	1	H2DL_STRCO
4	64	59.8	337	1	CT86_HUMAN
5	60	56.1	450	1	A2AA_HUMAN
6	60	56.1	721	1	FUE2_RAT
7	60	56.1	1005	1	EPAS5_RAT
8	59	55.1	129	1	COXB_HUMAN
9	59	55.1	877	1	EPAS5_MOUSE
10	57	53.3	129	1	COXB_RAT
11	57	53.3	870	1	SRC2_HUMAN
12	55.5	51.9	550	1	RLUB_XANAC
13	55	51.4	247	1	SPR1_HUMAN
14	55	51.4	1046	1	IF2_STRAW
15	54	50.5	452	1	A2A_BOVIN
16	54	50.5	1733	1	VNUA_PRVKA
17	53.5	50.0	627	1	SPD2_NEPCL
18	53	49.5	421	1	PO41_MOUSE
19	53	49.5	1336	1	NM54_HUMAN
20	53	49.5	2715	1	MLL4_HUMAN
21	52.5	49.1	782	1	L100_ADE12
22	52	48.6	314	1	PMXB_HUMAN
23	52	48.6	314	1	PMXB_MOUSE
24	52	48.6	325	1	CYT_MOUSE
25	52	48.6	416	1	KLFF_HUMAN
26	52	48.6	676	1	ICP0_HSVBK
27	52	48.6	676	1	ICP0_HSVBK
28	52	48.6	964	1	IF2_COREFL
29	52	48.6	1004	1	IF2_COREFL
30	52	48.6	1487	1	ICF4_HSVBK
31	52	48.6	1487	1	ICF4_HSVBK
32	51.5	48.1	1250	1	TP3A_DROME
33	51	47.7	199	1	TF15_HUMAN

34 51 47.7 327 1 FBRL\_GIALA Q24957 giardia lam  
35 51 47.7 423 1 PO41\_HUMAN Q01851 homo sapien  
36 51 47.7 707 1 PUB2\_HUMAN Q92945 homo sapien  
37 51 47.7 1033 1 IF2\_STRCO Q8CJG8 streptomyce  
38 51 47.7 1355 1 CA21\_RANCA Q42350 rana catesb  
39 51 47.7 1372 1 CA21\_MOUSE Q01149 mus musculu  
40 51 47.7 1372 1 CA21\_RAT P02466 rattus norv  
41 50.5 47.2 297 1 TLX3\_CHICK O93387 gallus gall  
42 50.5 47.2 384 1 GRP1\_PETHY P09789 petunia hyb  
43 50.5 47.2 450 1 A2AA\_MOUSE Q01338 mus musculu  
44 50.5 47.2 483 1 KG3A\_HUMAN P49840 homo sapien  
45 50.5 47.2 483 1 KG3A\_RAT P18265 rattus norv

#### ALIGNMENTS

RESULT 1  
CTG1\_HUMAN STANDARD; PRT; 180 AA.  
AC P78358;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 28-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cancer/testis antigen 1 (Autoimmunogenic cancer/testis antigen NY-ESO-1).  
GN CTAG1 OR CTAG;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A.  
RX MEDLINE=97203161; PubMed=9050879;  
RA Chen Y.-T., Scanlan M.J., Sabin U., Tuereci O., Gure A.O., Tsang S.,  
RA Williamson B., Stockert B., Ffirendschuh M., Old L.J.;  
RT "A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening."  
RT Proc. Natl. Acad. Sci. U.S.A. 94:1914-1918(1997).  
[2]  
SEQUENCE FROM N.A.  
RP TISSUE=Meianoma;  
RX MEDLINE=98289662; PubMed=9626360;  
RA Lethe B., Lucas S., Michaux L., de Smet C., Godelaine D., Serrano A.,  
RA de Plaen B., Boon T.;  
RT "LAGE-1, a new gene with tumor specificity."  
RL Int. J. Cancer 76:903-908(1998).  
[3]  
SEQUENCE FROM N.A.  
RX MEDLINE=98436682; PubMed=9759882;  
RA Wang R.-F., Johnston S.L., Zeng G., Topalian S.L.,  
RA Schwartzentruber D.J., Rosenberg S.A.;  
RT "A breast and melanoma-shared tumor antigen: T cell responses to antigenic peptides translated from different open reading frames."  
RL J. Immunol. 161:3596-3606(1998).  
CC -1- TISSUE SPECIFICITY: Expressed in testis and ovary and in a wide variety of cancers. Detected in uterine myometrium.  
CC -1- SIMILARITY: Belongs to the CTAG family.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; U87459; AAB49693.1; --  
DR EMBL; AJ003149; CAA05908.1; --  
DR EMBL; AF038567; AAD05202.1; --  
DR Genew; HGNC:2491; CTAG1.  
KW MW; 300156; --  
Transmembrane; Antigen.

```
FT DOMAIN 5 82 GLY-RICH.
FT TRANSEM 156 172 POTENTIAL.
SQ SEQUENCE 180 AA; 17992 MW; B122C5C2C8B8E1569 CRC64;

Query Match 100.0%; Score 107; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 2,9e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGRPGAGARASPGGAPR 20
   |||||
DB 43 RGRPGAGARASPGGAPR 62

RESULT 2
CTG2_HUMAN
ID CTG2_HUMAN STANDARD; PRT; 210 AA.
AC 075638; 075637;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Canser/testis antigen 2 (LAGE-1 protein).
GN CTAG2 OR LAGE1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS LAGE-1A AND LAGE-1B).
RC TISSUE=Melanoma;
RX MEDLINE=96289662; PubMed=9626360;
RA Lethe B., Lucas S., Michaux L., de Smet C., Godelaine D., Serrano A.,
RA de Plaan E., Boon T.;
RT "LAGE-1, a new gene with tumor specificity.";
RL Int. J. Cancer 76:903-908 (1998).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=LAGE-1B;
CC IsoId=075638-1; Sequence=Displayed;
CC Name=LAGE-1A;
CC IsoId=075638-2; Sequence=VSP 004301;
CC -1- TISSUE SPECIFICITY: Testis and very low level in placenta and in
CC some uterus samples. Observed in 25-50% of tumor samples of
CC melanomas, non-small-cell lung carcinomas, bladder, prostate and
CC head and neck cancers.
CC -1- DOMAIN: A transmembrane domain is present in isoform LAGE-1A.
CC -1- SIMILARITY: Belongs to the CTAG family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ223093; CAAL1117.1;
CC EMBL; AJ223093; CAAL1116.1;
CC EMBL; AJ223040; CAAL1043.1;
CC EMBL; AJ223041; CAAL1044.1;
CC Genew; HGNC:2492; CTAG2.
CC MIM; 300396;
CC Polymorphism; Alternative splicing; Transmembrane; Antigen.
FT DOMAIN 5 79 GLY-RICH.
FT DOVALN 183 188 POLY-PRO.
FT VARSPLIC 135 210
FT SEQRTGTPPPPPGAGQDGGCGVAFNMFSAPI -> IR
FT LTAADHRLQLSISCSLQQLSLMWITQCFLEVFQAQPSG
FT QRR (in isoform LAGE-1A).
FT /FTId=VSP_004301.
FT R -> Q.
FT /FTId=VAR_007855.
FT E -> Q.
FT /FTId=VAR_007856.
FT
```

```
FT VARIANT 138 138 W -> R.
FT /FTId=VAR_007857.
SQ SEQUENCE 210 AA; 21119 MW; 8B80E00A8558B8B CRC64;

Query Match 92.5%; Score 99; DB 1; Length 210;
Best Local Similarity 95.0%; Pred. No. 0.00024;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RGRPGAGARASPGGAPR 20
   |||||
DB 43 RGRPGAGARASPGGAPR 62

RESULT 3
HYDL_STRCO
ID HYDL_STRCO STANDARD; PRT; 627 AA.
AC P42534; Q952L7;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative polyketide hydroxylase (EC 1.14.13.-) (White ORF VIII).
GN SC05321 OR SC069.12C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21986410; PubMed=12000953;
RA Bentley S.D., Chater K.P., Harris D.E., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147 (2002).
RN [2]
RP SEQUENCE OF 1-255 FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=94075247; PubMed=8253693;
RA Blanco G., Pereda A., Brian P., Mendez C., Chater K.F., Salas J.A.;
RT "A hydroxylase-like gene product contributes to synthesis of a
RT polyketide spore pigment in Streptomyces halstedii.";
RL J. Bacteriol. 175:8043-8048 (1993).
CC -1- FUNCTION: Involved in developmentally regulated synthesis of a
CC compound biosynthetically related to polyketide antibiotics which
CC is essential for spore color in Streptococcus coelicolor.
CC -1- COFACTOR: FAD (by similarity).
CC -1- SIMILARITY: BELONGS TO THE PHEA/TFDB FAMILY OF FAD MONOOXYGENASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL939123; CAB45603.1;
CC EMBL; X74213; CAAS2289.1;
CC PIR; T35608; T35608.
CC InterPro; IPR008160; Collagen.
CC InterPro; IPR000733; Flav. monooxygenase.
CC InterPro; IPR002938; Mox. FAD binding.
CC InterPro; IPR003042; Rng. monooxygenase.
CC Pfam; PF01494; FAD binding 3; 1.
CC Pfam; PF01360; Monooxygenase; 1.
```

```

DR PRINTS; PR00420; RNMNOXGNASE.
KW Oxidoreductase; Flavoprotein; PAD; Complete proteome.
FT NP_BIND 22 51 FAD (POTENTIAL).
FT NP_BIND 309 319 FAD (POTENTIAL).
FT NP_BIND 60 60 R -> A (IN REF. 2).
FT CONFLICT 145 145 L -> LH (IN REF. 2).
FT CONFLICT 234 234 C -> S (IN REF. 2).
SQ SEQUENCE 627 AA; 64557 MW; 74684A2A9E9511C CRC64;

Query Match 62.6%; Score 67; DB 1; Length 627;
Best Local Similarity 75.0%; Pred. No. 1.7;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GPRGAGARASGPGGG 17
DQ 447 GPRGAGARASGPGGG 462

RESULT 4
CT86_HUMAN STANDARD; PRT; 337 AA.
AC Q9B219;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein c20orf86.
GN C20ORF86.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
RA Bailey J., Barlow K.P., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A.G., Cowile G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leheslahti M.H., Leverhwa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McKay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilmink L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
CC -1- SIMILARITY: Contains 2 ANK repeats.
CC -1- SIMILARITY: Contains 1 ubiquitin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; ALJ354776; CAC17565.2; -

```

```

DR HSP; P42771; 1DC2.
DR Genew; HGNC:16217; C20orf86.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00023; ank; 2.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 2.
DR PROSITE; PS00088; ANK REPEAT; 2.
DR PROSITE; PS02977; ANK REP REGION; 1.
DR PROSITE; PS02997; UBIQUITIN_1; FALSE_NEG.
DR PROSITE; PS00053; UBIQUITIN_2; 1.
KW Hypothetical protein; ANK repeat; Repeat; Polymorphism.
FT DOMAIN 88 164 UBIQUITIN-LIKE.
FT REPEAT 211 241 ANK 1.
FT REPEAT 244 273 ANK 2.
FT VARIANT 287 287 R -> C (in dbSNP:584855).
FT SEQUENCE 337 AA; 36714 MW; EC8BA4AD414756CB CRC64;

Query Match 59.8%; Score 64; DB 1; Length 337;
Best Local Similarity 65.0%; Pred. No. 2;
Matches 13; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 RGPRGAGARASGPGGGAPR 20
DQ 3 RAAGAGGARRAGPTGGASR 22

RESULT 5
A2AA_HUMAN STANDARD; PRT; 450 AA.
ID A2AA_HUMAN STANDARD; PRT; 450 AA.
AC P08913; Q9BZK1;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Alpha-2A adrenergic receptor (Alpha-2A adrenoceptor) (Alpha-2AAR subtype C10).
GN ADRA2A OR ADRA2R OR ADRAR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89308571; PubMed=2568356;
RA Fraser C.M., Arakawa S., McCombie W.R., Venter J.C.;
RT "Cloning, sequence analysis, and permanent expression of a human
RT alpha 2-adrenergic receptor in Chinese hamster ovary cells. Evidence
RT for independent pathways of receptor coupling to adenylylate cyclase
RT attenuation and activation.";
RL J. Biol. Chem. 264:11754-11761(1989).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Platelet;
RX MEDLINE=88042789; PubMed=2823383;
RA Koblika B.K., Matsui H., Koblika T.S., Yang-Feng T.L., Francke U.,
RA Caron M.G., Lefkowitz R.J., Regan J.W.;
RT "Cloning, sequencing, and expression of the gene coding for the human
RT platelet alpha 2-adrenergic receptor.";
RL Science 238:650-656(1987).
RN [3]
RP REVISIONS TO 333-365.
RX MEDLINE=91009167; PubMed=2170371;
RA Guyer C.A., Horstman D.A., Wilson A.L., Clark J.D., Kragoe E.J. Jr.,
RA Limbird L.E.;
RT "Cloning, sequencing, and expression of the gene encoding the porcine
RT alpha 2-adrenergic receptor. Allosteric modulation by Na+, H+, and
RT amiloride analogs.";
RL J. Biol. Chem. 265:17307-17317(1990).
RN [4]
RP SEQUENCE FROM N.A.
RA Castellano M., Giacche' M., Rossi F., Rivadossi F., Perani C.,
RA Beschi M., Agabiti Rosei E.;

```

RT "A search for genetic variability in the human alpha-2 adrenergic  
RT receptor on chromosome 10."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A., AND VARIANT LYS-251.  
RX MEDLINE=20556293; PubMed=10948191;  
RA Small K.M., Forbes S.L., Brown K.M., Liggett S.B.;  
RT "An Asn to Lys polymorphism in the third intracellular loop of the  
RT human alpha 2A-adrenergic receptor imparts enhanced agonist-promoted  
RT Gi coupling.";  
RL J. Biol. Chem. 275:38518-38523(2000).  
RN [6]  
RP SEQUENCE FROM N.A.  
RA Mao Z.-M., Tang K., Li B.-M., Jing N.-H.;  
RT "Cloning and expression of human alpha-2A adrenergic receptor in SY5Y  
RT cells.";  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RA Liu L., Yuan L.;  
RT "Human alpha-2A adrenergic receptor gene and the genotype of -1296  
RT nucleotide and motional sickness.";  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
RN [8]  
RP MUTAGENESIS OF PHE-412.  
RX MEDLINE=91332079; PubMed=1678390;  
RA Suryanarayana S., Daunt D.A., von Zastrow M., Kobilka B.K.;  
RT "A point mutation in the seventh hydrophobic domain of the alpha 2  
RT adrenergic receptor increases its affinity for a family of beta  
RT receptor antagonists.";  
RL J. Biol. Chem. 266:15488-15492(1991).  
RN [9]  
RP MUTAGENESIS OF ASPARTIC ACID AND SERINE RESIDUES.  
RX MEDLINE=91342598; PubMed=1678850;  
RA Wang C.-D., Buck M.A., Fraser C.W.;  
RT "Site-directed mutagenesis of alpha 2A-adrenergic receptors:  
RT Identification of amino acids involved in ligand binding and receptor  
RT activation by agonists.";  
RL Mol. Pharmacol. 40:168-179(1991).  
CC -!- FUNCTION: Alpha-2 adrenergic receptors mediate the catecholamine-  
CC induced inhibition of adenylate cyclase through the action of G  
CC proteins. The rank order of potency for agonists of this receptor  
CC is oxytocan > clonidine > epinephrine > norepinephrine >  
CC phenylephrine > dopamine > p-synephrine > p-tyramine > serotonin =  
CC p-octopamine. For antagonists, the rank order is yohimbine >  
CC phentolamine = mianserin > chlorpromazine = spiperone = prazosin  
CC > propranolol > alprenolol = pindolol.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M23533; AAA51665.1; -;  
CC EMBL; M18415; AAA51664.1; -;  
CC EMBL; AF262016; AAG00447.2; -;  
CC EMBL; AF281308; AAF91441.1; -;  
CC EMBL; AF316894; AAK01634.1; -;  
CC EMBL; AF284095; AAK26743.1; -;  
CC EMBL; AY032736; AAK51162.1; -;  
CC PIR; A34169; A34169.  
CC PDB; 1H09; 24-JUL-02.  
CC PDB; 1H0D; 24-JUL-02.  
CC PDB; 1H0F; 24-JUL-02.  
CC Genew; HGNC:281; ADRA2A.  
CC MIM; 104210; -;  
CC GO; GO:0005887; C:integral to plasma membrane; TAS.

DR GO; GO:0004938; F:alpha2-adrenergic receptor activity; TAS.  
DR GO; GO:0015459; F:potassium channel regulator activity; TAS.  
DR GO; GO:0030036; P:actin cytoskeleton organization and biogenesis; TAS.  
DR GO; GO:0000187; P:activation of MAPK; TAS.  
DR GO; GO:0006928; P:cell motility; TAS.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.  
DR GO; GO:0007194; P:negative regulation of adenylate cyclase ac. . .; TAS.  
DR GO; GO:0008284; P:positive regulation of cell proliferation; TAS.  
DR GO; GO:0007265; P:RAS protein signal transduction; TAS.  
DR GO; GO:0007266; P:Rho protein signal transduction; TAS.  
DR GO; GO:0007165; P:signal transduction; TAS.  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PRINTS; PR00237; GPCRHOPOPSN.  
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL\_1; 1.  
DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL\_2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
KW Multigene family; Phosphorylation; Lipoprotein; Palmitate;  
KW Polymorphism; 3D-structure.  
FT DOMAIN 1 33 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 34 59 1 (POTENTIAL).  
FT DOMAIN 60 70 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 71 96 2 (POTENTIAL).  
FT DOMAIN 97 106 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 107 129 3 (POTENTIAL).  
FT DOMAIN 130 149 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 150 173 4 (POTENTIAL).  
FT DOMAIN 174 192 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 193 217 5 (POTENTIAL).  
FT DOMAIN 218 374 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 375 399 6 (POTENTIAL).  
FT DOMAIN 400 406 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 407 430 7 (POTENTIAL).  
FT DOMAIN 431 450 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 14 14 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT DISULFID 106 188 BY SIMILARITY.  
FT LIPID 442 442 S-palmitoyl cysteine (by similarity).  
FT SITE 113 113 IMPLICATED IN LIGAND BINDING.  
FT SITE 200 200 IMPLICATED IN CATECHOL AGONIST BINDING  
AND RECEPTOR ACTIVATION.  
FT SITE 204 204 IMPLICATED IN CATECHOL AGONIST BINDING  
AND RECEPTOR ACTIVATION.  
FT VARIAT 251 N -> K (rare polymorphism; frequency  
in Caucasians 0.004 and in African-  
Americans 0.05; 40% increase in agonist-  
promoted Gi coupling; dbSNP:1800035).  
FT MUTAGEN 79 79 /FTID-VAR 014957.  
FT D->N: NO CHANGE IN BINDING AFFINITY.  
FT ELIMINATES GUANINE NUCLEOTIDE-SENSITIVE  
AGONIST BINDING.  
FT D->N: NO BINDING TO YOHIMBINE. INCREASE  
IN ADENYLATE CYCLASE ACTIVITY.  
FT D->N: LOWER AFFINITY FOR AGONISTS.  
FT ELIMINATES GUANINE NUCLEOTIDE-SENSITIVE  
AGONIST BINDING.  
FT S->A: LOWER AFFINITY FOR AGONISTS. NO  
CHANGE IN GUANINE NUCLEOTIDE-SENSITIVE  
AGONIST BINDING.  
FT S->A: LOWER AFFINITY FOR AGONISTS.  
FT REDUCED GUANINE NUCLEOTIDE-SENSITIVE  
AGONIST BINDING.  
FT F->N: 350-FOLD REDUCED AFFINITY FOR  
ALPHA-2 ANTAGONIST YOHIMBINE. 3000-FOLD  
INCREASE FOR BETA-ANTAGONIST ALPRENOLOL.  
FT A -> T (IN REF. 2).  
FT V -> C (IN REF. 2).  
FT PRRGPGATGIGTPAAGGEEVGAAGAKASWRGR -> RGAG  
RGRGSGRLQGRGASGLPRRRAGG (IN REF. 1  
AND 2).  
FT R -> L (IN REF. 2).  
FT CONFLICT 368 368  
FT SEQUENCE 450 AA; 48956 MW; A703CF262F04E8AC CRC64;

```

Query Match          56.1%; Score 60; DB 1; Length 450;
Best Local Similarity 63.6%; Pred. No. 7;
Matches 14; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY 1 RGRPGAGARAS--GPGGGAPR 20
    ||||| ||||| ||||| |||||
Db 313 RGRPGGKGKARASQVRKPGDSLPR 334

RESULT 6
FUB2 RAT
ID FUB2 RAT STANDARD; PRT; 721 AA.
AC Q99PF5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Far upstream element binding protein 2 (FUSE binding protein 2) (KH
DE type splicing regulatory protein) (KSRP) (MAP2 RNA trans-acting
DE protein 1) (MART1).
GN FUBP2 OR KHSRP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.; SEQUENCE OF 73-87; 89-108 AND 475-486, SUBCELLULAR
RP LOCATION, AND FUNCTION.
RC TISSUE=Brain;
RX MEDLINE=22246918; PubMed=12358751;
RA Rehbein M., Wege K., Buck F., Schweizer M., Richter D., Kindler S.;
RT "Molecular characterization of MART1, a protein interacting with the
RT dendritic targeting element of MAP2 mRNA."
RL J. Neurochem. 82:1039-1046(2002).
CC -!- FUNCTION: Part of a ternary complex that binds to the downstream
CC control sequence (DCS) of the pre-mRNA. Mediates exon inclusion in
CC transcripts that are subject to tissue-specific alternative
CC splicing. May interact with single-stranded DNA from the far-
CC upstream element (FUSE). May activate gene expression (by
CC similarity). Binds to the dendritic targeting element and may play
CC a role in mRNA trafficking.
CC -!- SUBUNIT: Part of a ternary complex containing FUBP2, PTBP1, PTBP2
CC and HNRPH1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear. A small proportion is also found in
CC the cytoplasm of neuronal cell bodies and dendrites.
CC -!- SIMILARITY: Contains 4 KH domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF308818; AAG59811.1; -.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH; 4.
DR SMART; SM00322; KH; 4.
DR PROSITE; PS50084; KH_TYPE_1; 4.
KW Transport; mRNA transport; mRNA processing; mRNA splicing;
KW Transcription regulation; Trans-acting factor; Nuclear protein;
KW DNA-binding; RNA-binding; Repeat.
FT DOMAIN 145 209
FT DOMAIN 234 300
FT DOMAIN 323 387
FT DOMAIN 425 492
FT DOMAIN 572 685
FT REPEAT 572 583
FT REPEAT 618 629
FT REPEAT 644 655
FT REPEAT 674 685
FT DOMAIN 7
FT GLY/PRO-RICH.

```

---

```

FT DOMAIN 69 498 GLY-RICH.
FT DOMAIN 499 613 ALA/GLY/PRO-RICH.
SQ SEQUENCE 721 AA; 74226 MW; 482C7A765C60EB4A CRC64;

Query Match          56.1%; Score 60; DB 1; Length 721;
Best Local Similarity 61.1%; Pred. No. 11;
Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 GPRGAGARASGPGGGAP 19
    ||||| ||||| ||||| |||||
Db 33 GPPGAGDRGGGPGGGP 50

RESULT 7
EPAS RAT
ID EPAS RAT STANDARD; PRT; 1005 AA.
AC P54757;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ephrin type-A receptor 5 precursor (EC 2.7.1.112) (Tyrosine-protein
DE kinase receptor EHK-1) (EPH homology kinase-1).
GN EPAS OR EKH1 OR EHK-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4 AND 5).
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=9406777; PubMed=7504232;
RA Maisongierre P.C., Barrezueta N.X., Yancopoulos G.D.;
RT "Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine
RT kinase family with distinctive structures and neuronal expression."
RL Oncogene 8:3277-3288(1993).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 6).
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=95206467; PubMed=7898646;
RA Taylor V., Miescher G.C., Pfarr S., Honegger P., Breitschopf H.,
RT Lassmann H., Steck A.J.;
RT "Expression and developmental regulation of Ehk-1, a neuronal
RT Elk-like receptor tyrosine kinase in brain."
RL Neuroscience 63:163-178(1994).
CC -!- FUNCTION: Receptor for members of the ephrin-A family. Binds to
CC ephrin-A1, -A2, -A3, -A4 and -A5.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=6;
CC Name=1;
CC IsoId=P54757-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P54757-2; Sequence=VSP_003001;
CC Name=3;
CC IsoId=P54757-3; Sequence=VSP_003002;
CC Name=4;
CC IsoId=P54757-4; Sequence=VSP_003002, VSP_003003;
CC Name=5;
CC IsoId=P54757-5; Sequence=VSP_003001, VSP_003002, VSP_003003;
CC Name=6;
CC IsoId=P54757-6; Sequence=VSP_003000, VSP_003002;
CC -!- TISSUE SPECIFICITY: ALMOST EXCLUSIVELY EXPRESSED IN THE NERVOUS
CC SYSTEM. PREDOMINANTLY EXPRESSED IN NEURONS.
CC -!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
CC -!- SIMILARITY: Contains 2 fibronectin type III domains.
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. Ephrin
CC receptor subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```



RA Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 32-44.  
 RC TISSUE=Liver;  
 RX MEDLINE=94147969; PubMed=8313870;  
 RA Hughes G.J., Frutiger S., Paquet N., Pasquali C., Sanchez J.-C.,  
 RA Tissot J.-D., Bairoch A., Appel R.D., Hochstrasser D.F.;  
 RT "Human liver protein map: update 1993";  
 RL Electrophoresis 14:1216-1222(1993).  
 CC -!- FUNCTION: This protein is one of the nuclear-coded polypeptide  
 CC chains of cytochrome c oxidase, the terminal oxidase in  
 CC mitochondrial electron transport.  
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycochrome c + O(2) = 4 ferricytochrome  
 CC c + 2 H(2)O.  
 CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.  
 CC -!- SIMILARITY: Belongs to the cytochrome c oxidase Vb family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M59250; AAA52060.1; -;  
 CC EMBL; M19961; AAA52061.1; -;  
 CC EMBL; BC062229; AAH06229.1; -;  
 CC FIRM; J03024; OTHUSB.  
 CC SWISS-2DPAGE; P10606; HUMAN.  
 CC Genew; HGNC:2269; COX5B.  
 CC MIM; 123866; -;  
 CC GO; GO:0004129; P:cytochrome-c oxidase activity; TAS.  
 CC GO; GO:0007585; P:respiratory gaseous exchange; TAS.  
 CC InterPro; IPR002124; COX5B.  
 CC Pfam; PF01215; COX5B; 1.  
 CC ProDom; PD007270; COX5B; 1.  
 CC PROSITE; PS00848; COX5B; 1.  
 CC Oxidoreductase; Inner membrane; Mitochondrion; Transit peptide;  
 KW Zinc.  
 FT TRANSIT 1 31 MITOCHONDRION.  
 FT CHAIN 32 129 CYTOCHROME C OXIDASE POLYPEPTIDE VB.  
 FT METAL 91 91 ZINC (POTENTIAL).  
 FT METAL 113 113 ZINC (POTENTIAL).  
 FT METAL 116 116 ZINC (POTENTIAL).  
 FT CONFLICT 109 109 E -> Q (IN REF. 1).  
 SQ SEQUENCES 129 AA; 13696 MW; 877BF4CD334AC931 CRC64;  
 Query Match 55.1%; Score 59; DB 1; Length 129;  
 Best Local Similarity 57.9%; Pred. No. 3;  
 Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
 QY 1 RGRPGAGAAASGPGGGAP 19  
 DB 20 RQPSGAARVSNASGGVVP 38  
 RESULT 9  
 ID EPAS MOUSE STANDARD; PRT; 877 AA.  
 AC Q60679;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-Oct-2001 (Rel. 40, Last annotation update)  
 DE Ephrin type-A receptor 5 precursor (EC 2.7.1.112) (Tyrosine-protein  
 DE kinase receptor EHK-1) (EPH homology kinase-1) (Brain-specific kinase)  
 DE (CEK-7).  
 GN EPHA5 OR EHK1 OR CEK7 OR BSK.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Brain;  
 RX MEDLINE=94194581; PubMed=8145300;  
 RA Zhou R., Copeland T.D., Kromer L.P., Schulz N.T.;  
 RT "Isolation and characterization of Bek, a growth factor receptor-like  
 RT tyrosine kinase associated with the limbic system";  
 RL J. Neurosci. Res. 37:129-143(1994).  
 CC -!- FUNCTION: Receptor for members of the ephrin-A family. Binds to  
 CC ephrin-A1, -A2, -A3, -A4 and -A5.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: BRAIN IN EMBRYO. NON NEURONAL EXPRESSION  
 CC DOMAINS INCLUDE THE ECTODERM OF THE BRANCHIAL ARCHES, THE ECTODERM  
 CC AND MESENCHYME SURROUNDING THE DORSAL ROOT GANGLIA, THE  
 CC INTERVERTEBRAL DISKS, MAXILLARY AND MANDIBULAR MESENCHYMAL  
 CC ELEMENTS AS WELL AS THE NASAL MESENCHYME AND ECTODERM. WITHIN THE  
 CC DEVELOPING NERVOUS SYSTEM, ITS EXPRESSION IS VERY DYNAMIC. BESIDES  
 CC THE MIDBRAIN IT IS ALSO EXPRESSED IN THE HYPOTHALAMUS, AND THE  
 CC NEUROHYPOPHYSIS.  
 CC -!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.  
 CC -!- SIMILARITY: Contains 1 fibronectin type III domain.  
 CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. Ephrin  
 CC receptor subfamily.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U07357; AAA17038.1; -;  
 CC FIRM; I48967; I48967.  
 CC HSSP; P29323; I48967.  
 CC MGD; MGI:99654; Ephas.  
 CC InterPro; IPR006209; EGF like.  
 CC InterPro; IPR001090; Ephrin receptor.  
 CC InterPro; IPR008957; FN III-like.  
 CC InterPro; IPR003961; FN III.  
 CC InterPro; IPR003962; FNIII subd.  
 CC InterPro; IPR008979; Gal\_bind like.  
 CC InterPro; IPR000719; Prot\_kinase.  
 CC InterPro; IPR001660; SAM.  
 CC InterPro; IPR001245; Tyr\_kinase.  
 CC InterPro; IPR008266; Tyr\_kinase AS.  
 CC InterPro; IPR001426; YKase\_receptorv.  
 CC Pfam; PF01404; EPH\_ibd; 1.  
 CC Pfam; PF00041; fn3; 1.  
 CC Pfam; PF00069; pkinase; 1.  
 CC Pfam; PF00536; SAM; 1.  
 CC PRINTS; PR00014; FNTYPEIII.  
 CC PRINTS; PR00109; TYRKINASE.  
 CC ProDom; PD001495; Ephrin\_receptor; 1.  
 CC ProDom; PD000001; Prot\_kinase; 1.  
 CC SMART; SMC0615; EPH\_ibd; 1.  
 CC SMART; SMC00600; FN3; 1.  
 CC SMART; SMC0454; SAM; 1.  
 CC SMART; SMC0219; TYRK; 1.  
 CC PROSITE; PS01186; EGF\_2; UNKNOWN 1.  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE; PS00790; RECEPTOR\_TYR\_KIN\_V1; 1.  
 CC PROSITE; PS00791; RECEPTOR\_TYR\_KIN\_V2; 1.  
 CC PROSITE; PS00105; SAM\_DOMAIN; 1.  
 CC Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 KW Receptor; Transmembrane; Glycoprotein; Signal.  
 PT SIGNAL 1 26 BY SIMILARITY.



Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences."  
Proc. Natl. Acad. Sci. U.S.A. 99:16895-16903(2002).  
CC -!- FUNCTION: Probable adhesion protein, which mediates homophilic and  
heterophilic interactions. In contrast to SCARF1, it poorly  
mediates the binding and degradation of acetylated low density  
lipoprotein (Ac-LDL) (By similarity).  
CC -!- SUBUNIT: Homophilic and heterophilic interaction via its  
extracellular domain. Interacts with SCARF1. The heterophilic  
interaction with SCARF1, which is stronger than the homophilic  
interaction with itself, is suppressed by the presence of SCARF1  
ligand such as Ac-LDL (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
CC -!- TISSUE SPECIFICITY: Predominantly expressed in endothelial cells.  
CC Expressed in heart, placenta, lung, kidney, spleen, small  
intestine and ovary.  
CC -!- SIMILARITY: Contains 7 EGF-like domains.  
CC -!- CAUTION: Ref.2 sequences differ from that shown due to  
frameshifts in positions 750, 751 and 768.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by not-for-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
EMBL; AF522196; AAN45861.1; -  
DR EMBL; BC000584; AAN00584.1; ALT\_FRAME.  
DR EMBL; BC009326; AAN09326.1; ALT\_FRAME.  
DR Genew; HGNC:19869; SCARF2.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR006210; IEGF.  
DR PRINTS; PRO0011; EGF\_LAMININ.  
DR SMART; SM00181; EGF; 7.  
DR SMART; SM00180; EGF\_Lam; 6.  
DR PROSITE; PS00022; EGF 1; 7.  
DR PROSITE; PS01186; EGF 2; 4.  
DR PROSITE; PS50026; EGF 3; 3.  
KW Cell adhesion; Receptor; Repeat; Signal; Transmembrane;  
KW EGF-like domain; Glycoprotein; Polymorphism.  
FT SIGNAL 1 43  
FT CHAIN 44 870 SCAVENGER RECEPTOR CLASS F MEMBER 2.  
FT DOMAIN 44 441 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 442 462  
FT DOMAIN 463 830 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 76 110 EGF-LIKE 1.  
FT DOMAIN 122 153 EGF-LIKE 2.  
FT DOMAIN 154 182 EGF-LIKE 3.  
FT DOMAIN 183 212 EGF-LIKE 4.  
FT DOMAIN 213 241 EGF-LIKE 5.  
FT DOMAIN 242 270 EGF-LIKE 6.  
FT DOMAIN 272 403 EGF-LIKE 7.  
FT DOMAIN 404 433 PRO-RICH.  
FT DISULFID 80 92 POTENTIAL.  
FT DISULFID 86 98 POTENTIAL.  
FT DISULFID 100 109 POTENTIAL.  
FT DISULFID 126 134 POTENTIAL.  
FT DISULFID 128 141 POTENTIAL.  
FT DISULFID 143 152 POTENTIAL.  
FT DISULFID 156 163 POTENTIAL.  
FT DISULFID 158 170 POTENTIAL.  
FT DISULFID 172 181 POTENTIAL.

FT DISULFID 185 193 POTENTIAL.  
FT DISULFID 187 200 POTENTIAL.  
FT DISULFID 202 211 POTENTIAL.  
FT DISULFID 215 222 POTENTIAL.  
FT DISULFID 217 229 POTENTIAL.  
FT DISULFID 231 240 POTENTIAL.  
FT DISULFID 244 251 POTENTIAL.  
FT DISULFID 246 258 POTENTIAL.  
FT DISULFID 260 269 POTENTIAL.  
FT DISULFID 376 384 POTENTIAL.  
FT DISULFID 379 391 POTENTIAL.  
FT DISULFID 393 402 POTENTIAL.  
FT CARBOHYD 83 83 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 365 365 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARIANT 777 777 D -> E (in dbSNP:759611).  
FT VARIANT 778 778 /FTID=VAR\_015148.  
FT VARIANT 778 778 V -> L (in dbSNP:759612).  
FT VARIANT 819 819 /FTID=VAR\_015149.  
FT VARIANT 837 837 A -> G (in dbSNP:874100).  
FT VARIANT 837 837 /FTID=VAR\_015150.  
FT VARIANT 837 837 /FTID=VAR\_015151.  
FT CONFLICT 474 478 MISSING (IN REF. 2).  
FT CONFLICT 626 641 ALYVARVARREARPARA -> GTRPTTWTTHSTAAS (IN  
REF. 2; AAN00584).  
SQ SEQUENCE 870 AA; 92479 MW; DCB735A50E6B9D1F CRC64;  
Query Match 53.3%; Score 57; DB 1; Length 870;  
Best Local Similarity 73.3%; Pred. No. 26;  
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 2 GPRGAGARASGPGG 16  
||||| ||| |||  
DB 5 GPRGAGPARRRGAGG 19  
RESULT 12  
RLUB XANAC  
ID RLUB XANAC STANDARD; PRT: 550 AA.  
AC Q8PK58;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Ribosomal large subunit pseudouridine synthase B (EC 4.2.1.70)  
DE (Pseudouridylylate synthase) (Uracil hydrolyase).  
GN RLUB OR XAC2318.  
OS Xanthomonas axonopodis (pv. citri).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
OC Xanthomonadaceae; Xanthomonas.  
OX NCBI\_TaxID=92829;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=306 / ATCC 13902 / XV 101;  
RX MEDLINE=2202145; PubMed=12024217;  
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Purlan L.R.,  
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,  
RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,  
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,  
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
RA Setubal J.C., Kitajima J.P.;  
RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
host specificities."

```

RL Nature 417:459-463(2002).
CC -|- FUNCTION: Responsible for synthesis of pseudouridine from
CC uracil-2605 in 23S ribosomal RNA (By similarity).
CC -|- CATALYTIC ACTIVITY: Uracil + D-ribose 5-phosphate = pseudouridine
CC 5'-phosphate + H(2)O.
CC -|- SIMILARITY: Belongs to the pseudouridine synthase rsuA family.
CC -|- SIMILARITY: Contains 1 S4 RNA-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB011865; AAM37171.1; -.
CC InterPro; IPR006145; Pseudou synth.
CC InterPro; IPR000743; Psi synth_RSU.
CC InterPro; IPR002942; S4 synth_RSU.
CC Pfam; PF00849; Pseudou synth_2; 1.
CC TIGRFAMs; TIGR00093; TIGR00093; 1.
CC PROSITE; PS01149; PSI_RSU; 1.
CC PROSITE; PS00889; S4; 1.
CC RNA processing; Lyase; RNA-binding; Complete proteome.
CC DOMAIN 28 97 S4 RNA-binding.
CC ACT SITE 132 132 BY SIMILARITY.
CC SEQUENCE 550 AA; 57729 MW; 548B3A448C8C04A2 CRC64;
FT
FT ACT SITE 132 132 BY SIMILARITY.
FT
FT SEQUENCE 550 AA; 57729 MW; 548B3A448C8C04A2 CRC64;
SQ
Query Match 51.9%; Score 55.5; DB 1; Length 550;
Best Local Similarity 61.9%; Pred. No. 25;
Matches 13; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 1 RGPGR-AGCAARASPGGCGAPR 20
Db 508 RGPGRPGCGARPGCGPGCGAPR 528

RESULT 13
SFRL1_HUMAN
ID SFRL1_HUMAN STANDARD; PRT; 247 AA.
AC Q07955; Q13809;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Splicing factor, arginine/serine-rich 1 (pre-mRNA splicing factor SF2,
DE P33 subunit) (Alternative splicing factor ASF-1).
GN SF21 OR SF2P33 OR ASF OR SF2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
PI 1
RN
SEQUENCE FROM N.A., AND SEQUENCE OF 143-160 AND 166-174.
RX MEDLINE=91309150; PubMed=1830244;
RA Krainer A.R., Mayeda A., Kozak D., Binns G.;
RT "Functional expression of cloned human splicing factor SF2: homology
RT to RNA-binding proteins, U1 70K, and Drosophila splicing
RT regulators.";
RL Cell 66:383-394(1991).
RN [1]
RN SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND ALTERNATIVE SPLICING.
RX MEDLINE=91309149; PubMed=1855257;
RA Ge H., Zuo P., Manley J.L.;
RT "Primary structure of the human splicing factor ASF reveals
RT similarities with Drosophila regulators.";
RL Cell 66:373-382(1991).
RN [3]
RN SEQUENCE FROM N.A. (ISOFORM ASF-1).
RX TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

```

```

RA Klauener R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murzy D.M., Sodargren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski W.I., Skalek U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RN SEQUENCE OF 122-139.
RX MEDLINE=92249775; PubMed=1577277;
RA Zahler A.M., Lane W.S., Stolk J.A., Roth M.B.;
RT "SR proteins: a conserved family of pre-mRNA splicing factors.";
RL Genes Dev. 6:837-847(1992).
RN [5]
RN INTERACTIONS IN SPLICEOSOME ASSEMBLY.
RX MEDLINE=94084782; PubMed=8261509;
RA Wu J.Y., Maniatis T.;
RT "Specific interactions between proteins implicated in splice site
RT selection and regulated alternative splicing.";
RL Cell 75:1061-1070(1993).
RN [6]
RN FUNCTION IN RECRUITMENT OF U1-70K TO PRE-MRNA.
RX MEDLINE=94187841; PubMed=8139654;
RA Kohtz J.D., Jamison S.F., Will C.L., Zuo P., Luhrmann R.,
RA Garcia-Blanco M.A., Manley J.L.;
RT "Protein-protein interactions and 5'-splice-site recognition in
RT mammalian mRNA precursors.";
RL Nature 368:119-124(1994).
RN [7]
RN RNA BINDING SPECIFICITY.
RX MEDLINE=95354672; PubMed=7543047;
RA Tacke R., Manley J.L.;
RT "The human splicing factors ASF/SF2 and SC35 possess distinct,
RT functionally significant RNA binding specificities.";
RL EMBO J. 14:3540-3551(1995).
RN [8]
RN MUTAGENESIS, AND CHARACTERIZATION OF FUNCTIONAL DOMAINS.
RX MEDLINE=94038956; PubMed=8223481;
RA Zuo P., Manley J.L.;
RT "Functional domains of the human splicing factor ASF/SF2.";
RL EMBO J. 12:4727-4737(1993).
RN [9]
RN RECOGNITION OF PRE-MRNA 5' SPLICING SITES.
RX MEDLINE=94211864; PubMed=7512732;
RA Zuo P., Manley J.L.;
RT "The human splicing factor ASF/SF2 can specifically recognize pre-mRNA
RT 5' splice sites.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:3363-3367(1994).
RN [10]
RN CHARACTERIZATION.
RX MEDLINE=92159037; PubMed=1741384;
RA Mayeda A., Zahler A.M., Krainer A.R., Roth M.B.;
RT "Two members of a conserved family of nuclear phosphoproteins are
RT involved in pre-mRNA splicing.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:1301-1304(1992).
RN [11]
RN INTERACTION WITH SAFB/SAFB1.
RX MEDLINE=98337913; PubMed=9671816;
RA Nayler O., Straetling W., Bourquin J.-P., Stagliar I., Lindemann L.,
RA Jaaper H., Hartmann A.M., Fackelmeyer F.O., Ullrich A., Stamm S.;
RT "SAF-B couples transcription and pre-mRNA splicing to SAR/MAR

```



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC EMBL; AF005031; BAC70262.1; -  
CC HANAP; MF\_00100; -; 1.  
CC InterPro; IPR001851; Bac\_inmem\_transp.  
CC InterPro; IPR004161; EFTU\_D2.  
CC InterPro; IPR000795; EF\_GTBind.  
CC InterPro; IPR003577; GTPase\_Ras.  
CC InterPro; IPR000178; IF2.  
CC InterPro; IPR006847; IF2\_N.  
CC InterPro; IPR002985; P-rich\_extensn.  
CC InterPro; IPR001806; Ras\_transfmrng.  
CC InterPro; IPR005225; Small\_GTP.  
CC InterPro; IPR009000; Translat\_factor.  
CC Pfam; PF00009; GTP\_EFTU; 1.  
CC Pfam; PF03144; GTP\_EFTU\_D2; 1.  
CC Pfam; PF04760; IF2\_N; 2.  
CC PRINTS; PR00315; ELONGATFNCT.  
CC PRINTS; PR01217; PRICHEXTENS.  
CC PRINTS; PR00449; RASTRNSFRMG.  
CC ProDom; PD186100; IF2; 1.  
CC SMART; SM00173; RAS; 1.  
CC TIGRfams; TIGR00487; IF-2; 1.  
CC TIGRfams; TIGR00231; small\_GTP; 1.  
CC PROSITE; PS01176; IF2; FALSE\_NEG.  
CC Initiation factor; Protein biosynthesis; GTP-binding;  
CC Complete proteome.  
CC DOMAIN 542 694 G-DOMAIN.  
FT NP\_BIND 548 555 GTP (BY SIMILARITY).  
FT NP\_BIND 598 602 GTP (BY SIMILARITY).  
FT NP\_BIND 652 655 GTP (BY SIMILARITY).  
SQ SEQUENCE 1046 AA; 106882 MW; 62B1E2DFA3A595 CRC64;

Query Match 51.4%; Score 55; DB 1; Length 1046;  
Best Local Similarity 60.0%; Pred. No. 50;  
Matches 12; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY 1 RGRGAGAGARASGPGGAPR 20  
DB 328 RGPGGGG--RPGGPGGGGR 345

RESULT 15  
ID A2AA\_BOVIN STANDARD; PRT; 452 AA.  
AC Q28838;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Alpha-2A adrenergic receptor (Alpha-2A adrenoceptor) (Alpha-2AAR)  
DE (Alpha-2D adrenergic receptor).  
GN ARA2A.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98111113; PubMed=9450652;  
RA Venkataraman V., Duda T., Sharma R.K.;  
RT "The bovine alpha 2D-adrenergic receptor gene: structure, expression  
RT in retina, and pharmacological characterization of the encoded  
RT receptor";  
RL Mol. Cell. Biochem. 177:113-123 (1997).  
RN [2]  
RP SEQUENCE OF 171-210 FROM N.A.

RX MEDLINE=94019366; PubMed=8412494;  
RA Blaxall H.S., Heck D.A., Bylund D.B.;  
RT "Molecular determinants of the alpha-2D adrenergic receptor subtype";  
RL Life Sci. 53:9-9(1993).  
CC !- FUNCTION: Alpha-2 adrenergic receptors mediate the catecholamine-  
CC induced inhibition of adenylylate cyclase through the action of G  
CC proteins.  
CC !- SUBCELLULAR LOCATION: Integral membrane protein.  
CC !- TISSUE SPECIFICITY: Retina, brain and olfactory lobe.  
CC !- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U79030; AAC24958.1; -  
CC EMBL; S66295; AAB28450.1; -  
CC PIR; I46958; I46958.  
CC HSSP; P29274; 1MMH.  
CC InterPro; IPR000276; GPCR\_Rhodopsn.  
CC Pfam; PF00001; 7tm\_1; 1.  
CC PRINTS; PR00237; GPCRHHODOPSN.  
CC PROSITE; PS00237; G-PROTEIN RECP F1\_1; 1.  
CC PROSITE; PS00262; G-PROTEIN RECP F1\_2; 1.  
CC G-protein coupled receptor; Transmembrane; Glycoprotein;  
CC Multigene family; Phosphorylation; Lipoprotein; Palmitate.  
FT DOMAIN 1 33 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 34 59 1 (POTENTIAL).  
FT DOMAIN 60 70 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 71 96 2 (POTENTIAL).  
FT DOMAIN 97 106 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 107 129 3 (POTENTIAL).  
FT DOMAIN 130 149 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 150 173 4 (POTENTIAL).  
FT DOMAIN 174 192 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 193 217 5 (POTENTIAL).  
FT DOMAIN 218 376 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 377 401 6 (POTENTIAL).  
FT DOMAIN 402 411 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 412 432 7 (POTENTIAL).  
FT DOMAIN 433 452 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 10 10 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 14 14 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT DISULFID 106 188 BY SIMILARITY.  
FT LIPID 444 444 S-palmitoyl cysteine (By similarity).  
SQ SEQUENCE 452 AA; 49252 MW; F10C1DD2860CD6F9 CRC64;

Query Match 50.5%; Score 54; DB 1; Length 452;  
Best Local Similarity 59.1%; Pred. No. 31;  
Matches 13; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

QY 1 RGRGAGAGARAS--GPGGAPR 20  
DB 314 RGPRAKAGARASQVKGDSLPR 335

Search completed: June 22, 2004, 08:51:57  
Job time : 16 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2004, 08:47:17 ; Search time 48 Seconds  
(without alignments)  
131.466 Million cell updates/sec

Title: US-09-529-206d-4\_COPY\_43\_62

Perfect score: 107

Sequence: 1 RGPAGAGARASGGGAPR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:

- 1: sp\_archaea:
- 2: sp\_bacteria:
- 3: sp\_fungi:
- 4: sp\_human:
- 5: sp\_invertebrate:
- 6: sp\_mammal:
- 7: sp\_mhc:
- 8: sp\_organelle:
- 9: sp\_phage:
- 10: sp\_plant:
- 11: sp\_rodent:
- 12: sp\_virus:
- 13: sp\_vertebrate:
- 14: sp\_unclassified:
- 15: sp\_rvirus:
- 16: sp\_bacteriaph:
- 17: sp\_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	142	Q9NY13	Q9NY13 homo sapien
2	99	92.5	180	Q9Y479	Q9Y479 homo sapien
3	99	92.5	210	Q9BU80	Q9BU80 homo sapien
4	99	92.5	210	Q9UJ89	Q9UJ89 homo sapien
5	62	57.9	301	Q8WS23	Q8WS23 nephila cla
6	62	57.9	550	Q8EP11	Q8EP11 toxoplasma
7	62	57.9	1039	Q8S0W4	Q8S0W4 oryza sativ
8	61.5	57.5	1953	Q9BIT7	Q9BIT7 nephila ina
9	60	56.1	465	Q8GTF8	Q8GTF8 homo sapien
10	60	56.1	926	Q8R554	Q8R554 mus musculu
11	59	55.1	420	Q8CBY9	Q8CBY9 mus musculu
12	59	55.1	538	Q8C278	Q8C278 mus musculu
13	59	55.1	581	Q8C276	Q8C276 mus musculu
14	58.5	54.7	563	Q9BIT5	Q9BIT5 nephila ina
15	58.5	54.7	566	Q19581	Q19581 caenorhabdi
16	58.5	54.7	692	Q19579	Q19579 caenorhabdi

#### ALIGNMENTS

##### RESULT 1

Q9NY13 ID Q9NY13 PRELIMINARY; PRT; 142 AA.

AC Q9NY13;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hypothetical protein (Fragment).  
GN LAGE-2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lethe B.G.  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ275978; CAB76945.1; -;  
KW Hypothetical protein.  
FT NON TER 1  
SQ SEQUENCE 142 AA; 13895 MW; 27EBE922AC4ACC7B CRC64;

Query Match 100.0%; Score 107; DB 4; Length 142;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGPAGAGARASGGGAPR 20

Db 17 RGPAGAGARASGGGAPR 36

##### RESULT 2

Q9Y479 ID Q9Y479 PRELIMINARY; PRT; 180 AA.

AC Q9Y479;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE LAGE-IS protein (Cancer/testis antigen 2).

```
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=93125550; PubMed=10399963;
RA Aarnoudse C.A., Van den Doel P.B., Heemskerk B., Schrier P.I.;
RT "Interleukin-2-induced, melanoma-specific T cells recognize CAMEL, an
unprecedented translation product of LAGE-1."
RL Int. J. Cancer 82:442-448(1999).
DR EMBL; AJ012835; CAA10194.1; -.
SQ SEQUENCE 180 AA; 18236 MW; 9077FAF953543A25 CRC64;

Query Match 92.5%; Score 99; DB 4; Length 180;
Best Local Similarity 95.0%; Pred. No. 0.00013;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RGRGAGAAASGPGGGAPR 20
Db 43 RGRGAGAAASGPGGGAPR 62

RESULT 3
Q9BU80 PRELIMINARY; PRT; 210 AA.
AC Q9BU80;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Cancer/testis antigen 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002833; AAB02833.1; -.
SQ SEQUENCE 210 AA; 21089 MW; 8FB5BF04FB04E8BE CRC64;

Query Match 92.5%; Score 99; DB 4; Length 210;
Best Local Similarity 95.0%; Pred. No. 0.00015;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RGRGAGAAASGPGGGAPR 20
Db 43 RGRGAGAAASGPGGGAPR 62

RESULT 4
Q9UJ89 PRELIMINARY; PRT; 210 AA.
AC Q9UJ89;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE LAGE-1L protein.
GN Homo sapiens (Human).
OS Homo sapiens (Human).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=99325550; PubMed=10399963;
RA Aarnoudse C.A., Van den Doel P.B., Heemskerk B., Schrier P.I.;
RT "Interleukin-2-induced, melanoma-specific T cells recognize CAMEL, an
unprecedented translation product of LAGE-1."
RL Int. J. Cancer 82:442-448(1999).
DR EMBL; AJ012835; CAA10194.1; -.
SQ SEQUENCE 210 AA; 21060 MW; 1DD081829735B60A CRC64;

Query Match 92.5%; Score 99; DB 4; Length 210;
Best Local Similarity 95.0%; Pred. No. 0.00015;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RGRGAGAAASGPGGGAPR 20
Db 43 RGRGAGAAASGPGGGAPR 62

RESULT 5
Q8WSZ3 PRELIMINARY; PRT; 301 AA.
AC Q8WSZ3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Dragline silk protein spidroin 2 (Fragment).
OS Nephila clavata (Goro spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Entelegynae; Araneioidea; Tetragnathidae; Nephila.
OX NCBI_TaxID=70342;
RN [1]
RP SEQUENCE FROM N.A.
RA Ma H.W., Zhang L.S., Lu Y.M., Liu Z.S., Zhang Y.J.;
RT "Amplification, Cloning and Sequence Analysis of Spider Dragline Silk
cDNA."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF41245; AAL32472.1; -.
FT NON_TER 1
FT NON_TER 301
SQ SEQUENCE 301 AA; 25396 MW; 8C69428B5684ED74 CRC64;

Query Match 57.9%; Score 62; DB 5; Length 301;
Best Local Similarity 66.7%; Pred. No. 6.1;
Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 GPRGAGAAASGPGGGAPR 19
Db 130 GPGGAAAAAAGPGGYGP 147

RESULT 6
Q86P11 PRELIMINARY; PRT; 550 AA.
AC Q86P11;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Jlp2.
OS Toxoplasma gondii.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae;
OC Toxoplasma.
OX NCBI_TaxID=5811;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RH;
RA Jiang L., Shu H., Luo S., Wu X., Cai L., Wang D., Zeng Q.;
RT "Immunological screening of Toxoplasma tachyzoite cDNA expression
libraries with serum from toxoplasma infected rats."
```

RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY208675; AAC03429.1; -

SQ SEQUENCE 550 AA; 57890 MW; 18C71B4974BF0F34 CRC64;

Query Match

Best Local Similarity 57.9%; Score 62; DB 5; Length 550;

Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GPRGAGAAARASGPGGAPR 20

Db 524 GPRAGAAQADGIGAGGPR 542

RESULT 7

ID Q8SOW4 PRELIMINARY; PRT; 1039 AA.

AC Q8SOW4;

DT 01-JUN-2002 (TRENBLrel. 21, Created)

DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)

DE 01-JUN-2002 (TRENBLrel. 24, Last annotation update)

DN OJ1014\_G12.10 protein.

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaceae; Oryza.

OX NCBI\_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;

RA sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, BAC

clone:OJ1014\_G12.10";

RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF003372; BAB89076.1; -

DR Gramene; Q8SOW4; -

DR InterPro; IPR007228; DUF390.

DR InterPro; IPR007321; Transposase\_28.

DR Pfam; PF04094; DUF390; 1.

DR Pfam; PF04195; Transposase\_28; 1.

SQ SEQUENCE 1039 AA; 112083 MW; A9A60FD3487B5DFD CRC64;

Query Match 57.9%; Score 62; DB 10; Length 1039;

Best Local Similarity 73.7%; Pred. No. 19;

Matches 14; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

Qy 3 PRGAGAAASG--PGGAP 19

Db 363 PRGGAGAAASRRPEGAP 381

RESULT 8

ID Q9BIT7

AC Q9BIT7 PRELIMINARY; PRT; 1953 AA.

DT 01-JUN-2001 (TRENBLrel. 17, Created)

DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)

DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)

DE Major ampullate spidroin 2-like protein (Fragment).

OS Nephila inaurata madagascariensis.

OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;

OC Araneomorphae; Entelegynae; Araneidae; Tetragnathidae; Nephila.

OX NCBI\_TaxID=115959;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21179804; PubMed=11283372;

RA Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;

RT "Extreme Diversity, Conservation, and Convergence of Spider Silk

Fibron Sequences";

RL Science 291:2603-2605(2001).

DR EMBL; AF350276; AAK30605.1; -

FT NON\_TER 1

FT NON\_TER 1953

SQ SEQUENCE 1953 AA; 159383 MW; C82B4DACD043C3BD CRC64;

Query Match

Best Local Similarity 57.5%; Score 61.5; DB 5; Length 1953;

Matches 13; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 1 RGPAGAAARASGPGGAP 19

Db 152 QGPRGAGAA--AAGPGGYGP 169

RESULT 9

ID Q86TH8

AC Q86TH8 PRELIMINARY; PRT; 465 AA.

DT 01-JUN-2003 (TRENBLrel. 24, Created)

DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)

DE Adrenergic, alpha-2A-, receptor.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Peripheral Nervous System;

RA Strausberg R.;

RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC050414; AAH50414.1; -

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0004872; F:receptor activity; IEA.

DR GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.

DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.

DR InterPro; IPR000276; GPCR\_Rhodopsn.

DR Pfam; PF00001; 7tm.1; 1.

DR PRINTS; PR00237; GPCRHOOPS.

DR PROSITE; PS00237; G\_PROTEIN\_RECPT\_FL\_1; 1.

DR PROSITE; PS0262; G\_PROTEIN\_RECPT\_FL\_2; 1.

KW Receptor.

SQ SEQUENCE 465 AA; 50646 MW; 585E576149BDB696 CRC64;

Query Match

Best Local Similarity 56.1%; Score 60; DB 4; Length 465;

Matches 14; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

Qy 1 RGPAGAAARAS--GPGGAPR 20

Db 328 RGPAGKARASQVKGDSLPR 349

RESULT 10

ID Q8R554

AC Q8R554 PRELIMINARY; PRT; 926 AA.

DT 01-JUN-2002 (TRENBLrel. 21, Created)

DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)

DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)

DE Cezanne 2 protein.

GN AJ430384 OR CEZANNE 2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=C57 BL/6;

RA Evans P.C., Coadwell W.J., Kilshaw P.J.;

RT "Isolation of a novel murine gene, Cezanne 2.;"

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ430384; CAD23048.1; -

DR MGD; MGI:2158505; AJ430384.

DR InterPro; IPR003323; OTU.

DR PROSITE; PS50802; OTU; 1.

RC	STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RD	MEDLINE=22354683; PubMed=12466851;
RE	The FANTOM Consortium.
RF	the RIKEN Genome Exploration Research Group Phase I & II Team;
RG	"Analysis of the mouse transcriptome based on functional annotation of
RH	60,770 full-length cDNAs";
RI	Nature 420:563-573(2002).
RJ	EMBL; AK089130; BAC40761.1; --
RK	GO; GO:0016020; C:membrane; IEA.
RL	GO; GO:0005524; F:ATP binding; IEA.
RM	GO; GO:0005003; P:ephra receptor activity; IEA.
RN	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
RO	GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . ; IEA.
RP	InterPro; IPR006209; EGF like.
RQ	InterPro; IPR001090; Ephrin receptor.
RR	InterPro; IPR003962; FN.III_subd.
RS	InterPro; IPR003961; FN.III.
RT	InterPro; IPR008957; FN.III-like.
RU	InterPro; IPR008979; Gal_bind_like.
RV	InterPro; IPR001426; YKase_receptorV.
RW	Pfam; PF01404; EPH_lbd; 1.
RX	Pfam; PF00041; fn3.1.
RY	PRINTS; PR00014; PNTYPE.III.
RZ	ProDom; PD001495; Ephrin_receptor; 1.
SA	SMART; SM00615; EPH_lbd; 1.
SB	SMART; SM00060; FN3.1.
SC	PROSITE; PS01186; EGF 2; 1.
SD	PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
SE	PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
SEQ	SEQUENCE 538 AA, 59823 MW, CP48C4P981452515 CRC64;

  

Query Match	55.1%;	Score 59;	DB 11;	Length 538;
Best Local Similarity	61.9%;	Pred. No. 24;		
Matches 13;	Conservative	0;	Mismatches	6;
			Indels	2;
			Gaps	1;

  

QY	2	GRPGACAARASGPGGG-APR 20
DB	5	GRPGAGHRRRTQGGGGDTPR 25

  

RESULT 13			
Q8C276	PRELIMINARY;	PRT;	681 AA.
AC	Q8C276;		
DT	01-MAR-2003 (TrEMBLrel. 23, Created)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Eph receptor A5.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RP	STRAIN=C57BL/6J; TISSUE=Olfactory brain;		
RC	MEDLINE=22354683; PubMed=12466851;		
RE	The FANTOM Consortium.		
RF	"Analysis of the mouse transcriptome based on functional annotation of		
RG	60,770 full-length cDNAs";		
RI	Nature 420:563-573(2002).		
RJ	EMBL; AK089143; BAC40764.1; --		
RK	GO; GO:0016020; C:membrane; IEA.		
RL	GO; GO:0005524; F:ATP binding; IEA.		
RM	GO; GO:0005003; P:ephra receptor activity; IEA.		
RO	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.		
RP	GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . ; IEA.		
RQ	InterPro; IPR006209; EGF like.		
RR	InterPro; IPR001090; Ephrin receptor.		
RS	InterPro; IPR003962; FN.III_subd.		
RT	InterPro; IPR003961; FN.III.		
RU	InterPro; IPR008957; FN.III-like.		
RV	InterPro; IPR008979; Gal_bind_like		

```
DR InterPro: IPR001426; YKase_receptorV.  
DR Pfam: PF01404; EPH_lbd; 1.  
DR Pfam: PF00041; fn3; 2.  
DR PRINTS: PRO0014; ENTPEI11.  
DR ProDom: PD061495; Ephrin_receptor; 1.  
DR SMART: SM00615; EPH_lbd; 1.  
DR SMART: SM00060; FN3; 2.  
DR PROSITE: PS01186; EGF_2; 1.  
DR PROSITE: PS00790; RECEPTOR_TYR_KIN_V.1; 1.  
DR PROSITE: PS00791; RECEPTOR_TYR_KIN_V.2; 1.  
SQ SEQUENCE 681 AA; 75336 MW; 92BB9187AB78BE722 CRC64;  
  
Query Match 55.1%; Score 59; DB 11; Length 681;  
Best Local Similarity 61.9%; Pred. No. 30;  
Matches 13; Conservative 0; Mismatches 6; Indels 2; Gaps 1;  
  
QY 2 GPRGAGARASGPGG--APR 20  
||||| ||||| ||||| |||||  
DB 5 GPRGAGHRTQGRGGDDTFR 25  
  
RESULT 14  
Q9BITS PRELIMINARY; PRT; 563 AA.  
AC Q9BITS  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
DE Major amputate epidroidin 2 (Fragment).  
GN MASP2.  
OS Nephila inaurata madagascariensis.  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;  
OC Araneomorphae; Entelegynae; Araneidae; Tetragnathidae; Nephila.  
OX NCBI_TaxID=115969;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21179804; PubMed=11283372;  
RT Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;  
RT "Extreme Diversity, Conservation, and Convergence of Spider Silk  
FT Fibroin Sequences";  
RL Science 291:2603-2605(2001).  
DR EMBL; AF350278; AAK30607.1; -.  
FT NON_TER 1  
FT NON_TER 563  
SQ SEQUENCE 563 AA; 48361 MW; 0AED885304B886B CRC64;  
  
Query Match 54.7%; Score 58.5; DB 5; Length 563;  
Best Local Similarity 59.1%; Pred. No. 29;  
Matches 13; Conservative 2; Mismatches 4; Indels 3; Gaps 1;  
  
QY 1 RGPRGAG--AARASGPGGGAP 19  
:||||| ||||| |||||  
DB 406 QGPGGAGSAAAAAAGPGGYGP 427  
  
RESULT 15  
Q19581 PRELIMINARY; PRT; 566 AA.  
AC Q19581  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE F18H3.3b protein.  
GN F18H3.3 OR F18H3.3B.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Coles L.;  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]
```

```
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
RT investigating biology.";  
RL Science 282:2012-2018(1998).  
DR EMBL; Z50110; CAA90446.1; -.  
DR PIR; T21096; T21096.  
DR HSSP; P11940; 1CVJ.  
DR WormPep; F18H3.3b; CE03230.  
DR GO; GO:0003676; P: nucleic acid binding; IEA.  
DR InterPro; IPR006515; PARP_1234.  
DR InterPro; IPR000504; RNA_rec_mot.  
DR Pfam; PF00076; rrm; 4.  
DR SMART; SM00360; RRM; 4.  
DR TIGRFAMs; TIGR01628; PABP-1234; 1.  
DR PROSITE; PS0102; RRM; 4.  
DR PROSITE; PS00030; RRM_RNP_1; 3.  
SQ SEQUENCE 566 AA; 62494 MW; A0BEF4AC7D4E5644 CRC64;  
  
Query Match 54.7%; Score 58.5; DB 5; Length 566;  
Best Local Similarity 65.0%; Pred. No. 29;  
Matches 13; Conservative 0; Mismatches 4; Indels 3; Gaps 1;  
  
QY 3 PRGAGARASGPG--CGAP 19  
||||| ||||| |||||  
DB 530 PRGAGPRVGGFGVQMGGAP 549  
  
Search completed: June 22, 2004, 08:53:23  
Job time : 50 secs
```